FIG.1A. DNA SEQUENCE OF HIGH MOLECULAR WEIGHT PROTEIN I (HMW1)

```
  1  ACAGCGTTCT CTTAATACTA GTACAAACCC ACAATAAAAT ATGACAAACA
 51  ACAATTACAA CACCTTTTTT GCAGTCTATA TGCAAATATT TTAAAAAATA
101  GTATAAATCC GCCATATAAA ATGGTATAAT CTTTCATCTT TCATCTTTCA
151  TCTTTCATCT TTCATCTTTC ATCTTTCATC TTTCATCTTT CATCTTTCAT
201  CTTTCATCTT TCATCTTTCA TCTTTCATCT TTCATCTTTC ACATGCCCTG
251  ATGAACCGAG GGAAGGGAGG GAGGGGCAAG AATGAAGAGG GAGCTGAACG
301  AACGCAAATG ATAAAGTAAT TTAATTGTTC AACTAACCTT AGGAGAAAAT
351  ATGAACAAGC TATATCGTCT CAAATTCAGC AAACGCCTGA ATGCTTTGGT
401  TGCTGTGTCT GAATTGGCAC GGGGTTGTGA CCATTCCACA GAAAAAGGCA
451  GCGAAAAACC TGCTCGCATG AAAGTGCGTC ACTTAGCGTT AAAGCCACTT
501  TCCGCTATGT TACTATCTTT AGGTGTAACA TCTATTCCAC AATCTGTTTT
551  AGCAAGCGGC TTACAAGGAA TGGATGTAGT ACACGGCACA GCCACTATGC
601  AAGTAGATGG TAATAAAACC ATTATCCGCA ACAGTGTTGA CGATATCATT
651  AATTGGAAAC AATTTAACAT CGACCAAAAT GAAATGGTGC AGTTTTTACA
701  AGAAAACAAC AACTCCGCCG TATTCAACCG TGTTACATCT AACCAAATCT
```

FIG.1B.

```
751   CCCAATTAAA AGGGATTTTA GATTCTAACG GACAAGTCTT TTTAATCAAC
801   CCAAATGGTA TCACAATAGG TAAAGACGCA ATTATTAACA CTAATGGCTT
851   TACGGCTTCT ACGCTAGACA TTTCTAACGA AAACATCAAG GCGCGTAATT
901   TCACCTTCGA GCAAACCAAA GATAAAGCGC TCGCTGAAAT TGTGAATCAC
951   GGTTTAATTA CTGTCGGTAA AGACGGCAGT GTAAATCTTA TTGGTGGCAA
1001  AGTGAAAAAC GAGGGTGTGA TTAGCGTAAA TGGTGGCAGC ATTTCTTTAC
1051  TCGCAGGGCA AAAAATCACC ATCAGCGATA TAATAAACCC AACCATTACT
1101  TACAGCATTG CCGCGCCTGA AAATGAAGCG GTCAATCTGG GCGATATTTT
1151  TGCCAAAGGC GGTAACATTA ATGTCCGTGC TGCCACTATT CGAAACCAAG
1201  GTAAACTTTC TGCTGATTCT GTAAGCAAAG ATAAAAGCGG CAATATTGTT
1251  CTTTCCGCCA AAGAGGGTGA AGCGGAAATT GGCGGTGTAA TTTCCGCTCA
1301  AAATCAGCAA GCTAAAGGCG GCAAGCTGAT GATTACAGGC GATAAAGTCA
1351  CATTAAAAAC AGGTGCAGTT ATCGACCTTT CAGGTAAAGA AGGGGGAGAA
1401  ACTTACCTTG GCGGTGACGA GCGCGGCGAA GGTAAAAAGG GCATTCAATT
1451  AGCAAAGAAA ACCTCTTTAG AAAAAGGCTC AACCATCAAT GTATCAGGCA
1501  AAGAAAAAGG CGGACGCGCT ATTGTGTGGG GCGATATTGC GTTAATTGAC
```

FIG.1C.

```
1551  GGCAATATTA ACGCTCAAGG TAGTGGTGAT ATCGCTAAAA CCGGTGGTTT
1601  TGTGGAGACG TCGGGGCATG ATTTATTCAT CAAAGACAAT GCAATTGTTG
1651  ACGCCAAAGA GTGGTTGTTA GACCCGGATA ATGTATCTAT TAATGCAGAA
1701  ACAGCAGGAC GCAGCAATAC TTCAGAAGAC GATGAATACA CGGGATCCGG
1751  GAATAGTGCC AGCACCCCAA AACGAAACAA AGAAAAGACA ACATTAACAA
1801  ACACAACTCT TGAGAGTATA CTAAAAAAAG GTACCTTTGT TAACATCACT
1851  GCTAATCAAC GCATCTATGT CAATAGCTCC ATTAATTTAT CCAATGGCAG
1901  CTTAACTCTT TGGAGTGAGG GTCGGAGCGG TGGCGGCGTT GAGATTAACA
1951  ACGATATTAC CACCGGTGAT GATACCAGAG GTGCAAACTT AACAATTTAC
2001  TCAGGCGGCT GGGTTGATGT TCATAAAAAT ATCTCACTCG GGGCGCAAGG
2051  TAACATAAAC ATTACAGCTA AACAAGATAT CGCCTTTGAG AAAGGAAGCA
2101  ACCAAGTCAT TACAGGTCAA GGGACTATTA CCTCAGGCAA TCAAAAAGGT
2151  TTTAGATTTA ATAATGTCTC TCTAAACGGC ACTGGCAGCG GACTGCAATT
2201  CACCACTAAA AGAACCAATA AATACGCTAT CACAAATAAA TTTGAAGGGA
2251  CTTTAAATAT TTCAGGGAAA GTGAACATCT CAATGGTTTT ACCTAAAAAT
2301  GAAAGTGGAT ATGATAAATT CAAAGGACGC ACTTACTGGA ATTTAACCTC
```

FIG.1D.

```
2351  CTTAAATGTT TCCGAGAGTG GCGAGTTTAA CCTCACTATT GACTCCAGAC
2401  GAAGCGATAG TGCAGGCACA CTTACCCAGC CTTATAATTT AAACGGTATA
2451  TCATTCAACA AAGACACTAC CTTTAATGTT GAACGAAATG CAAGAGTCAA
2501  CTTTGACATC AAGGCACCAA TAGGGATAAA TAAGTATTCT AGTTTGAATT
2551  ACGCATCATT TAATGGAAAC ATTTCAGTTT CGGGAGGGGG GAGTGTTGAT
2601  TTCACACTTC TCGCCTCATC CTCTAACGTC CAAACCCCCG GTGTAGTTAT
2651  AAATTCTAAA TACTTTAATG TTTCAACAGG GTCAAGTTTA AGATTTAAAA
2701  CTTCAGGCTC AACAAAAACT GGCTTCTCAA TAGAGAAAGA TTTAACTTTA
2751  AATGCCACCG GAGGCAACAT AACACTTTTG CAAGTTGAAG GCACCGATGG
2801  AATGATTGGT AAAGGCATTG TAGCCAAAAA AAACATAACC TTTGAAGGAG
2851  GTAACATCAC CTTTGGCTCC AGGAAAGCCG TAACAGAAAT CGAAGGCAAT
2901  GTTACTATCA ATAACAACGC TAACGTCACT CTTATCGGTT CGGATTTTGA
2951  CAACCATCAA AAACCTTTAA CTATTAAAAA AGATGTCATC ATTAATAGCG
3001  GCAACCTTAC CGCTGGAGGC AATATTGTCA ATATAGCCGG AAATCTTACC
3051  GTTGAAAGTA ACGCTAATTT CAAAGCTATC ACAAATTTCA CTTTTAATGT
3101  AGGCGGCTTG TTTGACAACA AAGGCAATTC AAATATTTCC ATTGCCAAAG
3151  GAGGGGCTCG CTTTAAAGAC ATTGATAATT CCAAGAATTT AAGCATCACC
```

FIG.1E.

```
3201  ACCAACTCCA GCTCCACTTA CCGCACTATT ATAAGCGGCA ATATAACCAA
3251  TAAAAACGGT GATTTAAATA TTACGAACGA AGGTAGTGAT ACTGAAATGC
3301  AAATTGGCGG CGATGTCTCG CAAAAAGAAG GTAATCTCAC GATTTCTTCT
3351  GACAAAATCA ATATTACCAA ACAGATAACA ATCAAGGCAG GTGTTGATGG
3401  GGAGAATTCC GATTCAGACG CGACAAACAA TGCCAATCTA ACCATTAAAA
3451  CCAAAGAATT GAAATTAACG CAAGACCTAA ATATTTCAGG TTTCAATAAA
3501  GCAGAGATTA CAGCTAAAGA TGGTAGTGAT TTAACTATTG GTAACACCAA
3551  TAGTGCTGAT GGTACTAATG CCAAAAAAGT AACCTTTAAC CAGGTTAAAG
3601  ATTCAAAAAT CTCTGCTGAC GGTCACAAGG TGACACTACA CAGCAAAGTG
3651  GAAACATCCG GTAGTAATAA CAACACTGAA GATAGCAGTG ACAATAATGC
3701  CGGCTTAACT ATCGATGCAA AAAATGTAAC AGTAAACAAC AATATTACTT
3751  CTCACAAAGC AGTGAGCATC TCTGCGACAA GTGGAGAAAT TACCACTAAA
3801  ACAGGTACAA CCATTAACGC AACCACTGGT AACGTGGAGA TAACCGCTCA
3851  AACAGGTAGT ATCCTAGGTG GAATTGAGTC CAGCTCTGGC TCTGTAACAC
3901  TTACTGCAAC CGAGGGCGCT CTTGCTGTAA GCAATATTTC GGGCAACACC
3951  GTTACTGTTA CTGCAAATAG CGGTGCATTA ACCACTTTGG CAGGCTCTAC
```

FIG.1F.

```
4001  AATTAAAGGA ACCGAGAGTG TAACCACTTC AAGTCAATCA GGCGATATCG
4051  GCGGTACGAT TTCTGGTGGC ACAGTAGAGG TTAAAGCAAC CGAAAGTTTA
4101  ACCACTCAAT CCAATTCAAA AATTAAAGCA ACAACAGGCG AGGCTAACGT
4151  AACAAGTGCA ACAGGTACAA TTGGTGGTAC GATTTCCGGT AATACGGTAA
4201  ATGTTACGGC AAACGCTGGC GATTTAACAG TTGGGAATGG CGCAGAAATT
4251  AATGCGACAG AAGGAGCTGC AACCTTAACT ACATCATCGG GCAAATTAAC
4301  TACCGAAGCT AGTTCACACA TTACTTCAGC CAAGGGTCAG GTAAATCTTT
4351  CAGCTCAGGA TGGTAGCGTT GCAGGAAGTA TTAATGCCGC CAATGTGACA
4401  CTAAATACTA CAGGCACTTT AACTACCGTG AAGGGTTCAA ACATTAATGC
4451  AACCAGCGGT ACCTTGGTTA TTAACGCAAA AGACGCTGAG CTAAATGGCG
4501  CAGCATTGGG TAACCACACA GTGGTAAATG CAACCAACGC AAATGGCTCC
4551  GGCAGCGTAA TCGCGACAAC CTCAAGCAGA GTGAACATCA CTGGGGATTT
4601  AATCACAATA AATGGATTAA ATATCATTTC AAAAAACGGT ATAAACACCG
4651  TACTGTTAAA AGGCGTTAAA ATTGATGTGA AATACATTCA ACCGGGTATA
4701  GCAAGCGTAG ATGAAGTAAT TGAAGCGAAA CGCATCCTTG AGAAGGTAAA
4751  AGATTTATCT GATGAAGAAA GAGAAGCGTT AGCTAAACTT GGAGTAAGTG
4801  CTGTACGTTT TATTGAGCCA AATAATACAA TTACAGTCGA TACACAAAAT
```

FIG.1G.

```
4851  GAATTTGCAA CCAGACCATT AAGTCGAATA GTGATTTCTG AAGGCAGGGC
4901  GTGTTTCTCA AACAGTGATG GCGCGACGGT GTGCGTTAAT ATCGCTGATA
4951  ACGGGCGGTA GCGGTCAGTA ATTGACAAGG TAGATTTCAT CCTGCAATGA
5001  AGTCATTTTA TTTTCGTATT ATTTACTGTG TGGGTTAAAG TTCAGTACGG
5051  GCTTTACCCA TCTTGTAAAA AATTACGGAG AATACAATAA AGTATTTTTA
5101  ACAGGTTATT ATTATG
```

FIG.2A. AMINO ACID SEQUENCE OF HIGH MOLECULAR WEIGHT PROTEIN I

```
  1  MNKIYRLKFS KRLNALVAVS ELARGCDHST EKGSEKPARM KVRHLALKPL
 51  SAMLLSLGVT SIPQSVLASG LQGMDVVHGT ATMQVDGNKT IIRNSVDAII
101  NWKQFNIDQN EMVQFLQENN NSAVFNRVTS NQISQLKGIL DSNGQVFLIN
151  PNGITIGKDA IINTNGFTAS TLDISNENIK ARNFTFEQTK DKALAEIVNH
201  GLITVGKDGS VNLIGGKVKN EGVISVNGGS ISLLAGQKIT ISDIINPTIT
251  YSIAAPENEA VNLGDIFAKG GNINVRAATI RNQGKLSADS VSKDKSGNIV
301  LSAKEGEAEI GGVISAQNQQ AKGGKLMITG DKVTLKTGAV IDLSGKEGGE
351  TYLGGDERGE GKNGIQLAKK TSLEKGSTIN VSGKEKGGRA IVWGDIALID
401  GNINAQGSGD IAKTGGFVET SGHDLFIKDN AIVDAKEWLL DFDNVSINAE
451  TAGRSNTSED DEYTGSGNSA STPKRNKEKT TLTNTTLESI LKKGTFVNIT
501  ANQRIYVNSS INLSNGSLTL WSEGRSGGGV EINNDITTGD DTRGANLTIY
551  SGGWVDVHKN ISLGAQGNIN ITAKQDIAFE KGSNQVITGQ GTITSGNQKG
601  FRFNNVSLNG TGSGLQFTTK RTNKYAITNK FEGTLNISGK VNISMVLPKN
651  ESGYDKFKGR TYWNLTSLNV SESGEFNLTI DSRGSDSAGT LTQPYNLNGI
701  SFNKDTTFNV ERNARVNFDI KAPIGINKYS SLNYASFNGN ISVSGGGSVD
```

FIG.2B.

```
 751  FTLLASSSNV QTPGVVINSK YFNVSTGSSL RFKTSGSTKT GFSIEKDLTL
 801  NATGGNITLL QVEGTDGMIG KGIVAKKNIT FEGGNITFGS RKAVTEIEGN
 851  VTINNNANVT LIGSDFDNHQ KPLTIKKDVI INSGNLTAGG NIVNIAGNLT
 901  VESNANFKAI TNFTFNVGGL FDNKGNSNIS IAKGGARFKD IDNSKNLSIT
 951  TNSSSTYRTI ISGNITNKNG DLNITNEGSD TEMQIGGDVS QKEGNLTISS
1001  DKINITKQIT IKAGVDGENS DSDATNNANL TIKTKELKLT QDLNISGFNK
1051  AEITAKDGSD LTIGNTNSAD GTNAKKVTFN QVKDSKISAD GHKVTLHSKV
1101  ETSGSNNNTE DSSDNNAGLT IDAKNVTVNN NITSHKAVSI SATSGEITTK
1151  TGTTINATTG NVEITAQTGS ILGGIESSSG SVTLTATEGA LAVSNISGNT
1201  VTVTANSGAL TTLAGSTIKG TESVTTSSQS GDIGGTISGG TVEVKATESL
1251  TTQSNSKIKA TTGEANVTSA TGTIGGTISG NTVNVTANAG DLTVGNGAEI
1301  NATEGAATLT TSSGKLTTEA SSHITSAKGQ VNLSAQDGSV AGSINAANVT
1351  LNTTGTLTTV KGSNINATSG TLVINAKDAE LNGAALGNHT VVNATNANGS
1401  GSVIATTSSR VNITGDLITI NGLNIISKNG INTVLLKGVK IDVKYIQPGI
1451  ASVDEVIEAK RILEKVKDLS DEEREALAKL GVSAVRFIEP NNTITVDTQN
1501  EFATRPLSRI VISEGRACFS NSDGATVCVN IADNGR
```

FIG.3A. DNA SEQUENCE OF HIGH MOLECULAR WEIGHT PROTEIN II (HMW2)

```
  1  TAAATATACA AGATAATAAA AATAAATCAA GATTTTTGTG ATGACAAACA
 51  ACAATTACAA CACCTTTTTT GCAGTCTATA TGCAAATATT TTAAAAAAAT
101  AGTATAAATC CGCCATATAA AATGGTATAA TCTTTCATCT TTCATCTTTA
151  ATCTTTCATC TTTCATCTTT CATCTTTCAT CTTTCATCTT TCATCTTTCA
201  TCTTTCATCT TTCATCTTTC ATCTTTCATC TTTCATCTTT CACATGAAAT
251  GATGAACCGA GGGAAGGGAG GGAGGGGCAA GAATGAAGAG GGAGCTGAAC
301  GAACGCAAAT GATAAAGTAA TTTAATTGTT CAACTAACCT TAGGAGAAAA
351  TATGAACAAG ATATATCGTC TCAAATTCAG CAAACGCCTG AATGCTTTGG
401  TTGCTGTGTC TGAATTGGCA CGGGGTTGTG ACCATTCCAC AGAAAAAGGC
451  TTCCGCTATG TTACTATCTT TAGGTGTAAC CACTTAGCGT TAAAGCCACT
501  TTCCGCTATG TTACTATCTT TAGGTGTAAC ATCTATTCCA CAATCTGTTT
551  TAGCAAGCGG CTTACAAGGA ATGGATGTAG TACACGGCAC AGCCACTATG
601  CAAGTAGATG GTAATAAAAC CATTATCCGC AACAGTGTTG ACGCTATCAT
651  TAATTGGAAA CAATTTAACA TCGACCAAAA TGAAATGGTG CAGTTTTTAC
701  AAGAAAACAA CAACTCCGCC GTATTCAACC GTGTTACATC TAACCAAATC
```

FIG.3B.

```
 751  TCCCAATTAA AAGGGATTTT AGATTCTAAC GGACAAGTCT TTTTAATCAA
 801  CCCAAATGGT ATCACAATAG GTAAAGACGC AATTATTAAC ACTAATGGCT
 851  TTACGGCTTC TACGCTAGAC ATTTCTAACG AAAACATCAA GGCGCGTAAT
 901  TTCACCTTCG AGCAAACCAA AGATAAAGCG CTCGCTGAAA TTGTGAATCA
 951  CGGTTTAATT ACTGTCGGTA AAGACGGCAG TGTAAATCTT ATTGGTGGCA
1001  AAGTGAAAAA CGAGGGTGTG ATTAGCGTAA ATGGTGGCAG CATTTCTTTA
1051  CTCGCAGGGC AAAAAATCAC CATCAGCGAT ATAATAAACC CAACCATTAC
1101  TTACAGCATT GCCGCGCCTG AAAATGAAGC GGTCAATCTG GGCGATATTT
1151  TTGCCAAAGG CGGTAACATT AATGTCCGTG CTGCCACTAT TCGAAACCAA
1201  GGTAAACTTT CTGCTGATTC TGTAAGCAAA GATAAAAGCG GCAATATTGT
1251  TCTTTCCGCC AAAGAGGGTG AAGCGGAAAT TGGCGGTGTA ATTTCCGCTC
1301  AAAATCAGCA AGCTAAAGGC GGCAAGCTGA TGATTACAGG CGATAAAGTC
1351  ACATTAAAAA CAGGTGCAGT TATCGACCTT TCAGGTAAAG AAGGGGGAGA
1401  AACTTACCTT GGCGGTGACG AGCGCGGCGA AGGTAAAAAC GGCATTCAAT
1451  TAGCAAAGAA AACCTCTTTA GAAAAAGGCT CAACCATCAA TGTATCAGGC
1501  AAAGAAAAAG GCGGACGCGC TATTGTGTGG GGCGATATTG CGTTAATTGA
```

FIG.3C.

```
1551  CGGCAATATT AACGCTCAAG GTAGTGGTGA TATCGCTAAA ACCGGTGGTT
1601  TTGTGGAGAC ATCGGGGCAT TATTTATCCA TTGACAGCAA TGCAATTGTT
1651  AAAACAAAAG AGTGGTTGCT AGACCCTGAT GATGTAACAA TTGAAGCCGA
1701  AGACCCCCTT CGCAATAATA CCGGTATAAA TGATGAATTC CCAACAGGCA
1751  CCGGTGAAGC AAGCGACCCT AAAAAAAATA GCGAACTCAA AACAACGCTA
1801  ACCAATACAA CTATTTCAAATTATCTGAAA AACGCCTGGA CAATGAATAT
1851  AACGGCATCA AGAAAACTTA CCGTTAATAG CTCAATCAAC ATCGGAAGCA
1901  ACTCCCACTT AATTCTCCAT AGTAAAGGTC AGCGTGGCGG AGGCGTTCAG
1951  ATTGATGGAG ATATTACTTC TAAAGGCGGA AATTTAACCA TTTATTCTGG
2001  CGGATGGGTT GATGTTCATA AAAATATTAC GCTTGATCAG GGTTTTTTAA
2051  ATATTACCGC CGCTTCCGTA GCTTTTGAAG GTGGAAATAA CAAAGCACGC
2101  GACGCGGCAA ATGCTAAAAT TGTCGCCCAG GGCACTGTAA CCATTACAGG
2151  AGAGGGAAAA GATTTCAGGG CTAACAACGT ATCTTTAAAC GGAACGGGTA
2201  AAGGTCTGAA TATCATTTCA TCAGTGAATA ATTTAACCCA CAATCTTAGT
2251  GGCACAATTA ACATATCTGG GAATATAACA ATTAACCAAA CTACGAGAAA
2301  GAACACCTCG TATTGGCAAA CCAGCCATGA TTCGCACTGG AACGTCAGTG
2351  CTCTTAATCT AGAGACAGGC GCAAATTTTA CCTTTATTAA ATACATTTCA
```

FIG.3D.

```
2401  AGCAATAGCA AAGGCTTAAC AACACAGTAT AGAAGCTCTG CAGGGGTGAA
2451  TTTTAACGGC GTAAATGGCA ACATGTCATT CAATCTCAAA GAAGGAGCGA
2501  AAGTTAATTT CAAATTAAAA CCAAACGAGA ACATGAACAC AAGCAAACCT
2551  TTACCAATTC GGTTTTTAGC CAATATCACA GCCACTGGTG GGGGCTCTGT
2601  TTTTTTTGAT ATATATGCCA ACCATTCTGG CAGAGGGGCT GAGTTAAAAA
2651  TGAGTGAAAT TAATATCTCT AACGGCGCTA ATTTTACCTT AAATTCCCAT
2701  GTTCGCGGCG ATGACGCTTT TAAAATCAAC AAAGACTTAA CCATAAATGC
2751  AACCAATTCA AATTTCAGCC TCAGACAGAC GAAAGATGAT TTTTATGACG
2801  GGTACGCACG CAATGCCATC AATTCAACCT ACAACATATC CATTCTGGGC
2851  GGTAATGTCA CCCTTGGTGG ACAAAACTCA AGCAGCAGCA TTACGGGGAA
2901  TATTACTATC GAGAAAGCAG CAAATGTTAC GCTAGAAGCC AATAACGCCC
2951  CTAATCAGCA AAACATAAGG GATAGAGTTA TAAAACTTGG CAGCTTGCTC
3001  GTTAATGGGA GTTTAAGTTT AACTGGCGAA AATGCAGATA TTAAAGGCAA
3051  TCTCACTATT TCAGAAAGCG CCACTTTTAA AGGAAAGACT AGAGATACCC
3101  TAAATATCAC CGGCAATTTT ACCAATAATG GCACTGCCGA AATTAATATA
3151  ACACAAGGAG TGGTAAAACT TGGCAATGTT ACCAATGATG GTGATTTAAA
```

FIG.3E.

```
3201  CATTACCACT CACGCTAAAC GCAACCAAAG AAGCATCATC GGCGGAGATA
3251  TAATCAACAA AAAAGGAAGC TTAAATATTA CAGACAGTAA TAATGATGCT
3301  GAAATCCAAA TTGGCGGCAA TATCTCGCAA AAAGAAGGCA ACCTCACGAT
3351  TTCTTCCGAT AAAATTAATA TCACCAAACA GATAACAATC AAAAAGGGTA
3401  TTGATGGAGA GGACTCTAGT TCAGATGCGA CAAGTAATGC CAACCTAACT
3451  ATTAAAACCA AAGAATTGAA ATTGACAGAA GACCTAAGTA TTTCAGGTTT
3501  CAATAAAGCA GAGATTACAG CCAAAGATGG TAGAGATTTA ACTATTGGCA
3551  ACAGTAATGA CGGTAACAGC GGTGCCGAAG CCAAAACAGT AACTTTTAAC
3601  AATGTTAAAG ATTCAAAAAT CTCTGCTGAC GGTCACAATG TGACACTAAA
3651  TAGCAAAGTG AAAACATCTA GCAGCAATGG CGGACGTGAA AGCAATAGCG
3701  ACAACGATAC CGGCTTAACT ATTACTGCAA AAAATGTAGA AGTAAACAAA
3751  GATATTACTT CTCTCAAAAC AGTAAATATC ACCGCGTCGG AAAAGGTTAC
3801  CACCACAGCA GGCTCGACCA TTAACGCAAC AAATGGCAAA GCAAGTATTA
3851  CAACCAAAAC AGGTGATATC AGCGGTACGA TTTCCGGTAA CACGGTAAGT
3901  GTTAGCGCGA CTGGTGATTT AACCACTAAA TCCGGCTCAA AAATTGAAGC
3951  GAAATCGGGT GAGGCTAATG TAACAAGTGC AACAGGTACA ATTGGCGGTA
```

FIG.3F.

```
4001  CAATTTCCGG TAATACGGTA AATGTTACGG CAAACGCTGG CGATTTAACA
4051  GTTGGGAATG GCGCAGAAAT TAATGCGACA GAAGGAGCTG CAACCTTAAC
4101  CGCAACAGGG AATACCTTGA CTACTGAAGC CGGTTCTAGC ATCACTTCAA
4151  CTAAGGGTCA GGTAGACCTC TTGGCTCAGA ATGGTAGCAT CGCAGGAAGC
4201  ATTAATGCTG CTAATGTGAC ATTAAATACT ACAGGCACCT TAACCACCGT
4251  GGCAGGCTCG GATATTAAAG CAACCAGCGG CACCTTGGTT ATTAACGCAA
4301  AAGATGCTAA GCTAAATGGT GATGCATCAG GTGATAGTAC AGAAGTGAAT
4351  GCAGTCAACG CAAGCGGCTC TGGTAGTGTG ACTGCGGCAA CCTCAAGCAG
4401  TGTGAATATC ACTGGGGATT TAAACACAGT AAATGGGTTA AATATCATTT
4451  CGAAAGATGG TAGAAACACT GTGCGCTTAA GAGGCAAGGA AATTGAGGTG
4501  AAATATATCC AGCCAGGTGT AGCAAGTGTA GAAGAAGTAA TTGAAGCGAA
4551  ACGCGTCCTT GAAAAAGTAA AAGATTTATC TGATGAAGAA AGAGAAACAT
4601  TAGCTAAACT TGGTGTAAGT GCTGTACGTT TTGTTGAGCC AAATAATACA
4651  ATTACAGTCA ATACACAAAA TGAATTTACA ACCAGACCGT CAAGTCAAGT
4701  GATAATTTCT GAAGGTAAGG CGTGTTTCTC AAGTGGTAAT GGCGCACGAG
4751  TATGTACCAA TGTTGCTGAC GATGGACAGC CGTAGTCAGT AATTGACAAG
4801  GTAGATTTCA TCCTGCAATG AAGTCATTTT ATTTTCGTAT TATTTACTGT
```

FIG.3G.

4851 GTGGGTTAAA GTTCAGTACG GGCTTTACCC ATCTTGTAAA AAATTACGGA
4901 GAATACAATA AAGTATTTTT AACAGGTTAT TATTATG

FIG.4A. AMINO ACID SEQUENCE OF HIGH MOLECULAR WEIGHT PROTEIN 2

```
  1  MNKIYRLKFS KRLNALVAVS ELARGCDHST EKGSEKPARM KVRHLALKPL
 51  SAMLLSLGVT SIPQSVLASG LQGMDVVHGT ATMQVDGNKT IIRNSVDAII
101  NWKQFNIDQN EMVQFLQENN NSAVFNRVTS NQISQLKGIL DSNGQVFLIN
151  PNGITIGKDA IINTNGFTAS TLDISNENIK ARNFTFEQTK DKALAEIVNH
201  GLITVGKDGS VNLIGGKVKN EGVISVNGGS ISLLAGQKIT ISDIINPTIT
251  YSIAAPENEA VNLGDIFAKG GNINVRAATI RNQGKLSADS VSKDKSGNIV
301  LSAKEGEAEI GGVISAQNQQ AKGGKLMITG DKVTLKTGAV IDLSGKEGGE
351  TYLGGDERGE GKNGIQLAKK TSLEKGSTIN VSGKEKGGRA IVWGDIALID
401  GNINAQGSGD IAKTGGFVET SGHDLFIKDN AIVDAKEWLL DFDNVSINAE
451  DPLRNNTGIN DEFPTGTGEA SDPKKNSELK TTLTNTTISN YLKNAWTMNI
501  TASRKLTVNS SINIGSNSHL ILHSKGQRGG GVQIDGDITS KGGNLTIYSG
551  GWVDVHKNIT LDQGFLNITA ASVAFEGGNN KARDAANAKI VAQGTVTITG
601  EGKDFRANNV SLNGTGKGLN IISSVNNLTH NLSGTINISG NITINQTTRK
651  NTSYWQTSHD SHWNVSALNL ETGANFTFIK YISSNSKGLT TQYRSSAGVN
701  FNGVNGNMSF NLKEGAKVNF KLKPNENMNT SKPLPIRFLA NITATGGGSV
```

FIG.4B.

```
 751  FFDIYANHSG RGAELKMSEI NISNGANFTL NSHVRGDDAF KINKDLTINA
 801  TNSNFSLRQT KDDFYDGYAR NAINSTYNIS ILGGNVTLGG QNSSSSITGN
 851  ITIEKAANVT LEANNAPNQQ NIRDRVIKLG SLLVNGSLSL TGENADIKGN
 901  LTISESATFK GKTRDTLNIT GNFTNNGTAE INITQGVVKL GNVTNDGDLN
 951  ITTHAKRNQR SIIGGDIINK KGSLNITDSN NDAEIQIGGN ISQKEGNLTI
1001  SSDKINITKQ ITIKKGIDGE DSSSDATSNA NLTIKTKELK LTEDLSISGF
1051  NKAEITAKDG RDLTIGNSND GNSGAEAKTV TFNNVKDSKI SADGHNVTLN
1101  SKVKTSSSNG GRESNSDNDT GLTITAKNVE VNKDITSLKT VNITASEKVT
1151  TTAGSTINAT NGKASITTKT GDISGTISGN TVSVSATVDL TTKSGSKIEA
1201  KSGEANVTSA TGTIGGTISG NTVNVTANAG DLTVGNGAEI NATEGAATLT
1251  ATGNTLTTEA GSSITSTKGQ VDLLAQNGSI AGSINAANVT LNTTGTLTTV
1301  AGSDIKATSG TLVINAKDAK LNGDASGDST EVNAVNASGS GSVTAATSSS
1351  VNITGDLNTV NGLNIISKDG RNTVRLRGKE IEVKYIQPGV ASVEEVIEAK
1401  RVLEKVKDLS DEERETLAKL GVSAVRFVEP NNTITVNTQN EFTTRPSSQV
1451  IISEGKACFS SGNGARVCTN VADDGQP
```

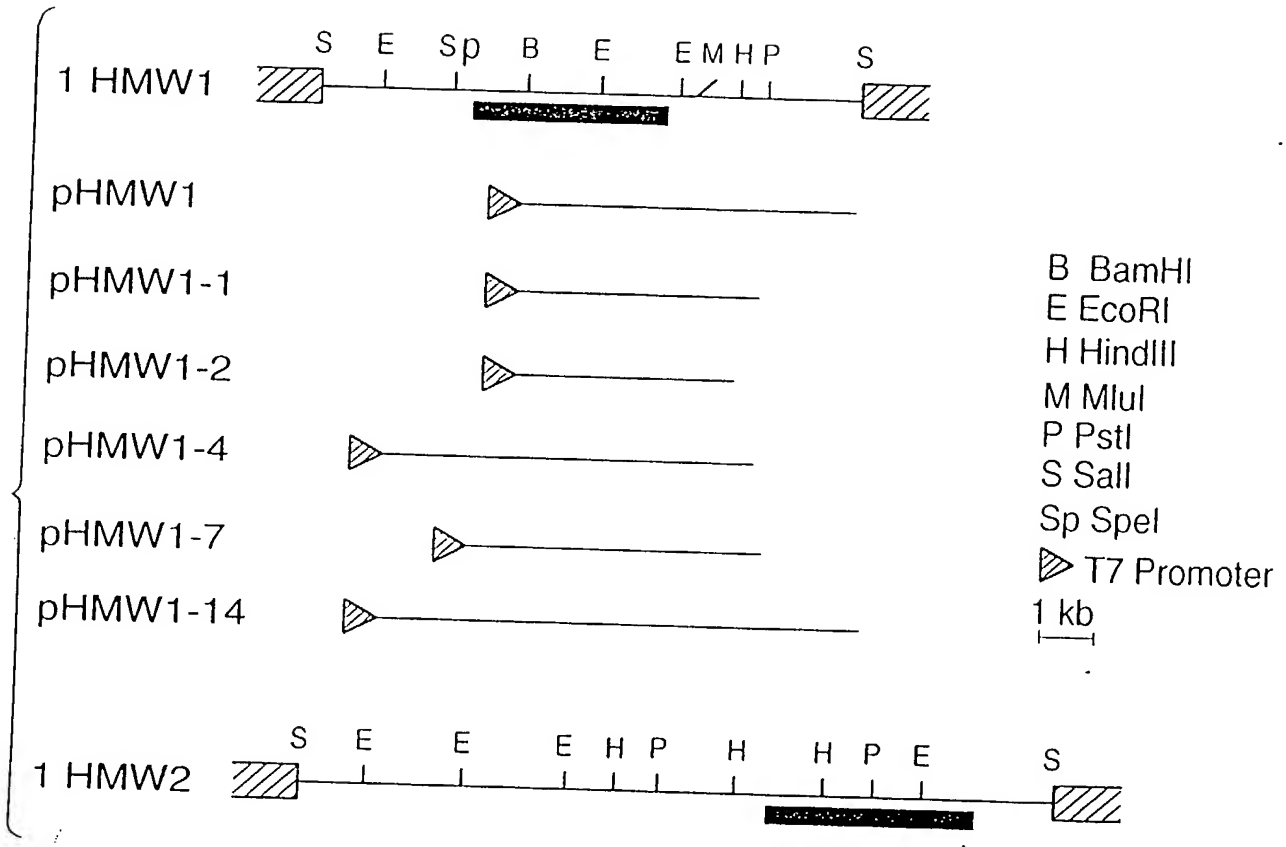

S E Sp B E E M H P S
1 HMW1
pHMW1
pHMW1-1
pHMW1-2
pHMW1-4
pHMW1-7
pHMW1-14
B BamHI
E EcoRI
H HindIII
M MluI
P PstI
S SalI
Sp SpeI
T7 Promoter
1 kb
S E E E H P H H P E S
1 HMW2

FIG.5A.

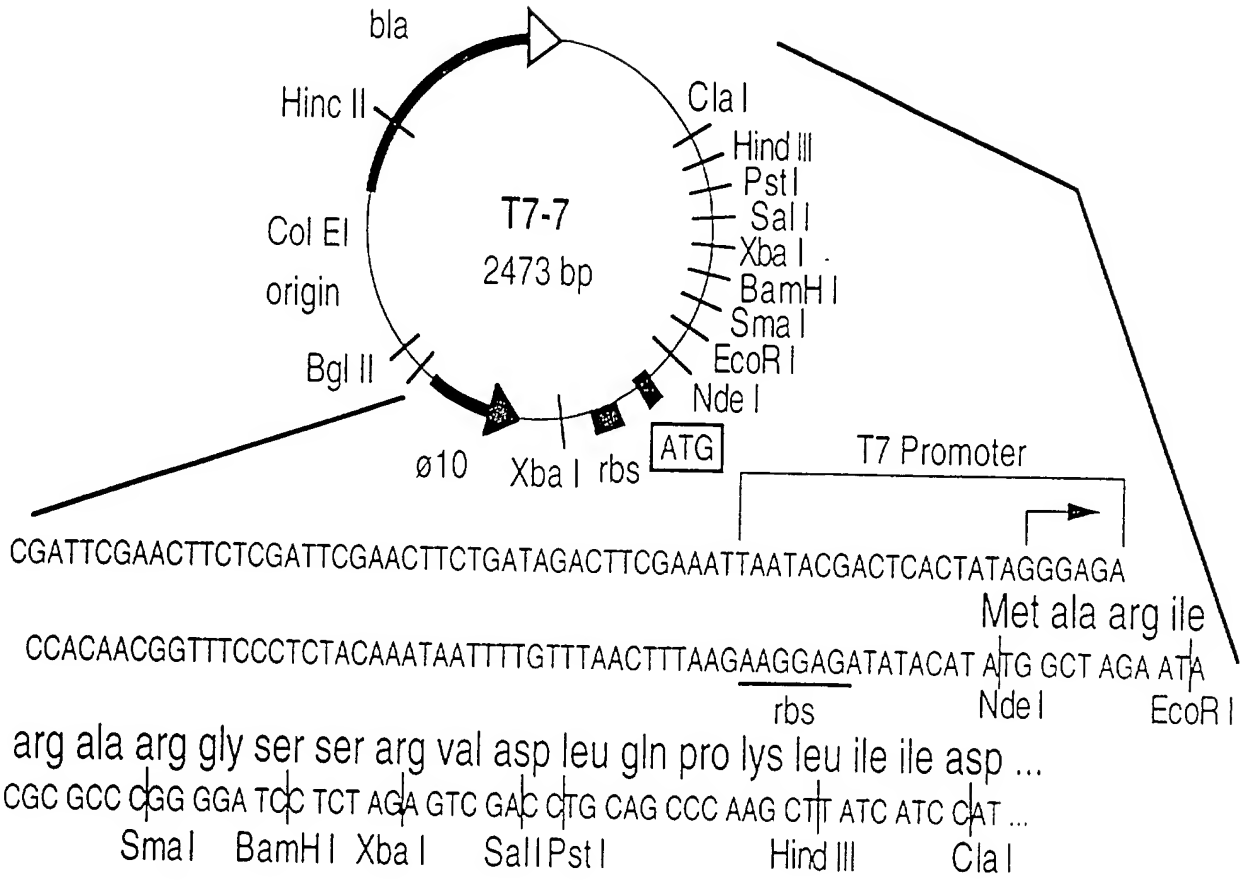


FIG.5B.

(A) Partial restriction maps of representative HMW1 and HMW2 recombinant phage and of HMW1 plasmid subclones. The shaded boxes indicate the locations of the structural genes. In the recombinant phage, transcription proceeds from left to right for the HMW1 gene and from right to left for the HMW2 gene. The methods used for construction of the plasmids shown are described in the text. (B) Restriction map of the T7 expression vector pT7-7. This vector contains the T7 RNA polymerase promoter ϕ10, a ribosome - binding site (rbs), and the translational start site for the T7 gene 10 protein upstream from a multiple cloning site (37).

FIG.6A.

```
   1  ACAGCGTTCT CTTAATACTA GTACAAACCC ACAATAAAAT ATGACAAACA
  51  ACAATTACAA CACCTTTTTT GCAGTCTATA TGCAAATATT TTAAAAAATA
 101  GTATAAATCC GCCATATAAA ATGGTATAAT CTTTCATCTT TCATCTTTCA
 151  TCTTTCATCT TTCATCTTTC ATCTTTCATC TTTCATCTTT CATCTTTCAT
 201  CTTTCATCTT TCATCTTTCA TCTTTCATCT TTCATCTTTC ACATGAAATG
 251  ATGAACCGAG GGAAGGGAGG GAGGGGCAAG AATGAAGAGG GAGCTGAACG
 301  AACGCAAATG ATAAAGTAAT TTAATTGTTC AACTAACCTT AGGAGAAAAT
 351  ATGAACAAGA TATATCGTCT CAAATTCAGC AAACGCCTGA ATGCTTTGGT
 401  TGCTGTGTCT GAATTGGCAC GGGGTTGTGA CCATTCCACA GAAAAAGGCA
 451  GCGAAAAACC TGCTCGCATG AAAGTGCGTC ACTTAGCGTT AAAGCCACTT
 501  TCCGCTATGT TACTATCTTT AGGTGTAACA TCTATTCCAC AATCTGTTTT
 551  AGCAAGCGGC TTACAAGGAA TGGATGTAGT ACACGGCACA GCCACTATGC
 601  AAGTAGATGG TAATAAAACC ATTATCCGCA ACAGTGTTGA CGCTATCATT
 651  AATTGGAAAC AATTTAACAT CGACCAAAAT GAAATGGTGC AGTTTTTACA
 701  AGAAAACAAC AACTCCGCCG TATTCAACCG TGTTACATCT AACCAAATCT
 751  CCCAATTAAA AGGGATTTTA GATTCTAACG GACAAGTCTT TTTAATCAAC
```

FIG.6B.

```
 801  CCAAATGGTA TCACAATAGG TAAAGACGCA ATTATTAACA CTAATGGCTT
 851  TACGGCTTCT ACGCTAGACA TTTCTAACGA AAACATCAAG GCGCGTAATT
 901  TCACCTTCGA GCAAACCAAA GATAAAGCGC TCGCTGAAAT TGTGAATCAC
 951  GGTTTAATTA CTGTCGGTAA AGACGGCAGT GTAAATCTTA TTGGTGGCAA
1001  AGTGAAAAAC GAGGGTGTGA TTAGCGTAAA TGGTGGCAGC ATTTCTTTAC
1051  TCGCAGGGCA AAAAATCACC ATCAGCGATA TAATAAACCC AACCATTACT
1101  TACAGCATTG CCGCGCCTGA AAATGAAGCG GTCAATCTGG GCGATATTTT
1151  TGCCAAAGGC GGTAACATTA ATGTCCGTGC TGCCACTATT CGAAACCAAG
1251  CTTTCCGCCA AAGAGGGTGA AGCGGAAATT GGCGGTGTAA TTTCCGCTCA
1301  AAATCAGCAA GCTAAAGGCG GCAAGCTGAT GATTACAGGC GATAAAGTCA
1351  CATTAAAAAC AGGTGCAGTT ATCGACCTTT CAGGTAAAGA AGGGGGAGAA
1401  ACTTACCTTG GCGGTGACGA GCGCGGCGAA GGTAAAAACG GCATTCAATT
1451  AGCAAAGAAA ACCTCTTTAG AAAAAGGCTC AACCATCAAT GTATCAGGCA
1501  AAGAAAAAGG CGGACGCGCT ATTGTGTGGG GCGATATTGC GTTAATTGAC
1551  GGCAATATTA ACGCTCAAGG TAGTGGTGAT ATCGCTAAAA CCGGTGGTTT
1601  TGTGGAGACG TCGGGGCATG ATTTATTCAT CAAAGACAAT GCAATTGTTG
```

FIG.6C.

```
1651  ACGCCAAAGA GTGGTTGTTA GACCCGGATA ATGTATCTAT TAATGCAGAA
1701  ACAGCAGGAC GCAGCAATAC TTCAGAAGAC GATGAATACA CGGGATCCGG
1751  GAATAGTGCC AGCACCCCAA AACGAAACAA AGAAAAGACA ACATTAACAA
1801  ACACAACTCT TGAGAGTATA CTAAAAAAAG GTACCTTTGT TAACATCACT
1851  GCTAATCAAC GCATCTATGT CAATAGCTCC ATTAATTTAT CCAATGGCAG
1901  CTTAACTCTT TGGAGTGAGG GTCGGAGCGG TGGCGGCGTT GAGATTAACA
1951  ACGATATTAC CACCGGTGAT GATACCAGAG GTGCAAACTT AACAATTTAC
2001  TCAGGCGGCT GGGTTGATGT TCATAAAAAT ATCTCACTCG GGGCGCAAGG
2051  TAACATAAAC ATTACAGCTA AACAAGATAT CGCCTTTGAG AAAGGAAGCA
2101  ACCAAGTCAT TACAGGTCAA GGGACTATTA CCTCAGGCAA TCAAAAAGGT
2151  TTTAGATTTA ATAATGTCTC TCTAAACGGC ACTGGCAGCG GACTGCAATT
2201  CACCACTAAA AGAACCAATA AATACGCTAT CACAAATAAA TTTGAAGGGA
2251  CTTTAAATAT TTCAGGGAAA GTGAACATCT CAATGGTTTT ACCTAAAAAT
2301  GAAAGTGGAT ATGATAAATT CAAAGGACGC ACTTACTGGA ATTTAACCTC
2351  GAAAGTGGAT ATGATAAATT CAAAGGACGC CCTCACTATT GACTCCAGAG
2401  GAAGCGATAG TGCAGGCACA CTTACCCAGC CTTATAATTT AAACGGTATA
2451  TCATTCAACA AAGACACTAC CTTTAATGTT GAACGAAATG CAAGAGTCAA
```

FIG.6D.

```
2501  CTTTGACATC AAGGCACCAA TAGGGATAAA TAAGTATTCT AGTTTGAATT
2551  ACGCATCATT TAATGGAAAC ATTTCAGTTT CGGGAGGGGG GAGTGTTGAT
2601  TTCACACTTC TCGCCTCATC CTCTAACGTC CAAACCCCCG GTGTAGTTAT
2651  AAATTCTAAA TACTTTAATG TTTCAACAGG GTCAAGTTTA AGATTTAAAA
2701  CTTCAGGCTC AACAAAAACT GGCTTCTCAA TAGAGAAAGA TTTAACTTTA
2751  AATGCCACCG GAGGCAACAT AACACTTTTG CAAGTTGAAG GCACCGATGG
2801  AATGATTGGT AAAGGCATTG TAGCCAAAAA AAACATAACC TTTGAAGGAG
2851  GTAAGATGAG GTTTGGCTCC AGGAAAGCCG TAACAGAAAT CGAAGGCAAT
2901  GTTACTATCA ATAACAACGC TAACGTCACT CTTATCGGTT CGGATTTTGA
2951  CAACCATCAA AAACCTTTAA CTATTAAAAA AGATGTCATC ATTAATAGCG
3001  GCAACCTTAC CGCTGGAGGC AATATTGTCA ATATAGCCGG AAATCTTACC
3051  GTTGAAAGTA ACGCTAATTT CAAAGCTATC ACAAATTTCA CTTTTAATGT
3101  AGGCGGCTTG TTTGACAACA AAGGCAATTC AAATATTTCC ATTGCCAAAG
3151  GAGGGGCTCG CTTTAAAGAC ATTGATAATT CCAAGAATTT AAGCATCACC
3201  ACCAACTCCA GCTCCACTTA CCGCACTATT ATAAGCGGCA ATATAACCAA
3251  TAAAAACGGT GATTTAAATA TTACGAACGA AGGTAGTGAT ACTGAAATGC
```

FIG.6E.

```
3301  AAATTGGCGG CGATGTCTCG CAAAAAGAAG GTAATCTCAC GATTTCTTCT
3351  GACAAAATCA ATATTACCAA ACAGATAACA ATCAAGGCAG GTGTTGATGG
3401  GGAGAATTCC GATTCAGACG CGACAAACAA TGCCAATCTA ACCATTAAAA
3451  CCAAAGAATT GAAATTAACG CAAGACCTAA ATATTTCAGG TTTCAATAAA
3501  GCAGAGATTA CAGCTAAAGA TGGTAGTGAT TTAACTATTG GTAACACCAA
3551  TAGTGCTGAT GGTACTAATG CCAAAAAAGT AACCTTTAAC CAGGTTAAAG
3601  ATTCAAAAAT CTCTGCTGAC GGTCACAAGG TGACACTACA CAGCAAAGTG
3651  GAAACATCCG GTAGTAATAA CAACACTGAA GATAGCAGTG ACAATAATGC
3701  CGGCTTAACT ATCGATGCAA AAAATGTAAC AGTAAACAAC AATATTACTT
3751  CTCACAAAGC AGTGAGCATC TCTGCGACAA GTGGAGAAAT TACCACTAAA
3801  ACAGGTACAA CCATTAACGC AACCACTGGT AACGTGGAGA TAACCGCTCA
3851  AACAGGTAGT ATCCTAGGTG GAATTGAGTC CAGCTCTGGC TCTGTAACAC
3901  TTACTGCAAC CGAGGGCGCT CTTGCTGTAA GCAATATTTC GGGCAACACC
3951  GTTACTGTTA CTGCAAATAG CGGTGCATTA ACCACTTTGG CAGGCTCTAC
4001  AATTAAAGGA ACCGAGAGTG TAACCACTTC AAGTCAATCA GGCGATATCG
4051  GCGGTACGAT TTCTGGTGGC ACAGTAGAGG TTAAAGCAAC CGAAAGTTTA
```

FIG.6F.

```
4101  ACCACTCAAT CCAATTCAAA AATTAAAGCA ACAACAGGCG AGGCTAACGT
4151  AACAAGTGCA ACAGGTACAA TTGGTGGTAC GATTTCCGGT AATACGGTAA
4201  ATGTTACGGC AAACGCTGGC GATTTAACAG TTGGGAATGG CGCAGAAATT
4251  AATGCGACAG AAGGAGCTGC AACCTTAACT ACATCATCGG GCAAATTAAC
4301  TACCGAAGCT AGTTCACACA TTACTTCAGC CAAGGGTCAG GTAAATCTTT
4351  CAGCTCAGGA TGGTAGCGTT GCAGGAAGTA TTAATGCCGC CAATGTGACA
4401  CTAAATACTA CAGGCACTTT AACTACCGTG AAGGGTTCAA ACATTAATGC
4451  AACCAGCGGT ACCTTGGTTA TTAACGCAAA AGACGCTGAG CTAAATGGCG
4501  CAGCATTGGG TAACCACACA GTGGTAAATG CAACCAACGC AAATGGCTCC
4551  GGCAGCGTAA TCGCGACAAC CTCAAGCAGA GTGAACATCA CTGGGGATTT
4601  AATCACAATA AATGGATTAA ATATCATTTC AAAAAACGGT ATAAACACCG
4651  TACTGTTAAA AGGCGTTAAA ATTGATGTGA AATACATTCA ACCGGGTATA
4701  GCAAGCGTAG ATGAAGTAAT TGAAGCGAAA CGCATCCTTG AGAAGGTAAA
4751  AGATTTATCT GATGAAGAAA GAGAAGCGTT AGCTAAACTT GGCGTAAGTG
4801  CTGTACGTTT TATTGAGCCA AATAATACAA TTACAGTCGA TACACAAAAT
4851  GAATTTGCAA CCAGACCATT AAGTCGAATA GTGATTTCTG AAGGCAGGGC
4901  GTGTTTCTCA AACAGTGATG GCGCGACGGT GTGCGTTAAT ATCGCTGATA
```

FIG.6G.

```
4951  ACGGGCGGTA GCGGTCAGTA ATTGACAAGG TAGATTTCAT CCTGCAATGA
5001  AGTCATTTTA TTTTCGTATT ATTTACTGTG TGGGTTAAAG TTCAGTACGG
5051  GCTTTACCCA TCTTGTAAAA AATTACGGAG AATACAATAA AGTATTTTTA
5101  ACAGGTTATT ATTATGAAAA ATATAAAAAG CAGATTAAAA CTCAGTGCAA
5151  TATCAGTATT GCTTGGCCTG GCTTCTTCAT CATTGTATGC AGAAGAAGCG
5201  TTTTTAGTAA AAGGCTTTCA GTTATCTGGT GCACTTGAAA CTTTAAGTGA
5251  AGACGCCCAA CTGTCTGTAG CAAAATCTTT ATCTAAATAC CAAGGCTCGC
5301  AAACTTTAAC AAACCTAAAA ACAGCACAGC TTGAATTACA GGCTGTGCTA
5351  GATAAGATTG AGCCAAATAA GTTTGATGTG ATATTGCCAC AACAAACCAT
5401  TACGGATGGC AATATTATGT TTGAGCTAGT CTCGAAATCA GCCGCAGAAA
5451  GCCAAGTTTT TTATAAGGCG AGCCAGGGTT ATAGTGAAGA AAATATCGCT
5501  CGTAGCCTGC CATCTTTGAA ACAAGGAAAA GTGTATGAAG ATGGTCGTCA
5551  GTGGTTCGAT TTGCGTGAAT TCAATATGGC AAAAGAAAAT CCACTTAAAG
5601  TCACTCGCGT GCATTACGAG TTAAACCCTA AAAACAAAAC CTCTGATTTG
5651  GTAGTTGCAG GTTTTTCGCC TTTTGGCAAA ACGCGTAGCT TTGTTTCCTA
5701  TGATAATTTC GGCGCAAGGG AGTTTAACTA TCAACGTGTA AGTCTAGGTT
```

FIG.6H.

```
5751  TTGTAAATGC CAATTTGACC GGACATGATG ATGTATTAAA TCTAAACGCA
5801  TTGACCAATG TAAAAGCACC ATCAAAATCT TATGCGGTAG GCATAGGATA
5851  TACTTATCCG TTTTATGATA AACACCAATC CTTAAGTCTT TATACCAGCA
5901  TGAGTTATGC TGATTCTAAT GATATCGACG GCTTACCAAG TGCGATTAAT
5951  CGTAAATTAT CAAAAGGTCA ATCTATCTCT GCGAATCTGA AATGGAGTTA
6001  TTATCTCCCG ACATTTAACC TTGGAATGGA AGACCAGTTT AAAATTAATT
6051  TAGGCTACAA CTACCGCCAT ATTAATCAAA CATCCGAGTT AAACACCCTG
6101  GGTGCAACGA AGAAAAAATT TGCAGTATCA GGCGTAAGTG CAGGCATTGA
6151  TGGACATATC CAATTTACCC CTAAAACAAT CTTTAATATT GATTTAACTC
6201  ATCATTATTA CGCGAGTAAA TTACCAGGCT CTTTTGGAAT GGAGCGCATT
6251  GGCGAAACAT TTAATCGCAG CTATCACATT AGCACAGCCA GTTTAGGGTT
6301  GAGTCAAGAG TTTGCTCAAG GTTGGCATTT TAGCAGTCAA TTATCGGGTC
6351  AGTTTACTCT ACAAGATATA AGTAGCATAG ATTTATTCTC TGTAACAGGT
6401  ACTTATGGCG TCAGAGGCTT TAAATACGGC GGTGCAAGTG GTGAGCGCGG
6451  TCTTGTATGG CGTAATGAAT TAAGTATGCC AAAATACACC CGCTTTCAAA
6501  TCAGCCCTTA TGCGTTTTAT GATGCAGGTC AGTTCCGTTA TAATAGCGAA
6551  AATGCTAAAA CTTACGGCGA AGATATGCAC ACGGTATCCT CTGCGGGTTT
```

FIG.6I.

```
6601  AGGCATTAAA ACCTCTCCTA CACAAAACTT AAGCTTAGAT GCTTTTGTTG
6651  CTCGTCGCTT TGCAAATGCC AATAGTGACA ATTTGAATGG CAACAAAAAA
6701  CGCACAAGCT CACCTACAAC CTTCTGGGGT AGATTAACAT TCAGTTTCTA
6751  ACCCTGAAAT TTAATCAACT GGTAAGCGTT CCGCCTACCA GTTTATAACT
6801  ATATGCTTTA CCCGCCAATT TACAGTCTAT ACGCAACCCT GTTTTCATCC
6851  TTATATATCA AACAAACTAA GCAAACCAAG CAAACCAAGC AAACCAAGCA
6901  AACCAAGCAA ACCAAGCAAA CCAAGCAAAC CAAGCAAACC AAGCAAACCA
6951  AGCAAACCAA GCAAACCAAG CAAACCAAGC AAACCAAGCA ATGCTAAAAA
7001  ACAATTTATA TGATAAACTA AAACATACTC CATACCATGG CAATACAAGG
7051  GATTTAATAA TATGACAAAA GAAAATTTAC AAAGTGTTCC ACAAAATACG
7101  ACCGCTTCAC TTGTAGAATC AAACAACGAC CAAACTTCCC TGCAAATACT
7151  TAAACAACCA CCCAAACCCA ACCTATTACG CCTGGAACAA CATGTCGCCA
7201  AAAAAGATTA TGAGCTTGCT TGCCGCGAAT TAATGGCGAT TTTGGAAAAA
7251  ATGGACGCTA ATTTTGGAGG CGTTCACGAT ATTGAATTTG ACGCACCTGC
7301  TCAGCTGGCA TATCTACCCG AAAAACTACT AATTCATTTT GCCACTCGTC
7351  TCGCTAATGC AATTACAACA CTCTTTTCCG ACCCCGAATT GGCAATTTCC
```

FIG.6J.

```
7401  GAAGAAGGGG CATTAAAGAT GATTAGCCTG CAACGCTGGT TGACGCTGAT
7451  TTTTGCCTCT TCCCCCTACG TTAACGCAGA CCATATTCTC AATAAATATA
7501  ATATCAACCC AGATTCCGAA GGTGGCTTTC ATTTAGCAAC AGACAACTCT
7551  TCTATTGCTA AATTCTGTAT TTTTTACTTA CCCGAATCCA ATGTCAATAT
7601  GAGTTTAGAT GCGTTATGGG CAGGGAATCA ACAACTTTGT GCTTCATTGT
7651  GTTTTGCGTT GCAGTCTTCA CGTTTTATTG GTACTGCATC TGCGTTTCAT
7701  AAAAGAGCGG TGGTTTTACA GTGGTTTCCT AAAAAACTCG CCGAAATTGC
7751  TAATTTAGAT GAATTGCCTG CAAATATCCT TCATGATGTA TATATGCACT
7801  GCAGTTATGA TTTAGCAAAA AACAAGCACG ATGTTAAGCG TCCATTAAAC
7851  GAACTTGTCC GCAAGCATAT CCTCACGCAA GGATGGCAAG ACCGCTACCT
7901  TTACACCTTA GGTAAAAAGG ACGGCAAACC TGTGATGATG GTACTGCTTG
7951  AACATTTTAA TTCGGGACAT TCGATTTATC GCACGCATTC AACTTCAATG
8001  ATTGCTGCTC GAGAAAAATT CTATTTAGTC GGCTTAGGCC ATGAGGGCGT
8051  TGATAACATA GGTCGAGAAG TGTTTGACGA GTTCTTTGAA ATCAGTAGCA
8101  ATAATATAAT GGAGAGACTG TTTTTTATCC GTAAACAGTG CGAAACTTTC
8151  CAACCCGCAG TGTTCTATAT GCCAAGCATT GGCATGGATA TTACCACGAT
```

FIG.6K.

```
8201  TTTTGTGAGC AACACTCGGC TTGCCCCTAT TCAAGCTGTA GCCTTGGGTC
8251  ATCCTGCCAC TACGCATTCT GAATTTATTG ATTATGTCAT CGTAGAAGAT
8301  GATTATGTGG GCAGTGAAGA TTGTTTAGC  GAAACCCTTT TACGCTTACC
8351  CAAAGATGCC CTACCTTATG TACCATCTGC ACTCGCCCCA CAAAAAGTGG
8401  ATTATGTACT CAGGGAAAAC CCTGAAGTAG TCAATATCGG TATTGCCGCT
8451  ACCACAATGA AATTAAACCC TGAATTTTTG CTAACATTGC AAGAAATCAG
8501  AGATAAAGCT AAAGTCAAAA TACATTTTCA TTTCGCACTT GGACAATCAA
8551  CAGGCTTGAC ACACCCTTAT GTCAAATGGT TTATCGAAAG CTATTTAGGT
8601  GACGATGCCA CTGCACATCC CCACGCACCT TATCACGATT ATCTGGCAAT
8651  ATTGCGTGAT TGCGATATGC TACTAAATCC GTTTCCTTTC GGTAATACTA
8701  ACGGCATAAT TGATATGGTT ACATTAGGTT TAGTTGGTGT ATGCAAAACG
8751  GGGGATGAAG TACATGAACA TATTGATGAA GGTCTGTTTA AACGCTTAGG
8801  ACTACCAGAA TGGCTGATAG CCGACACACG AGAAACATAT ATTGAATGTG
8851  CTTTGCGTCT AGCAGAAAAC CATCAAGAAC GCCTTGAACT CCGTCGTTAC
8901  ATCATAGAAA ACAACGGCTT ACAAAAGCTT TTTACAGGCG ACCCTCGTCC
8951  ATTGGGCAAA ATACTGCTTA AGAAAACAAA TGAATGGAAG CGGAAGCACT
9001  TGAGTAAAAA ATAACGGTTT TTTAAAGTAA AAGTGCGGTT AATTTTCAAA
```

FIG.6L.

```
9051  GCGTTTTAAA AACCTCTCAA AAATCAACCG CACTTTTATC TTTATAACGC
9101  TCCCGCGCGC TGACAGTTTA TCTCTTTCTT AAAATACCCA TAAAATTGTG
9151  GCAATAGTTG GGTAATCAAA TTCAATTGTT GATACGGCAA ACTAAAGACG
9201  GCGCGTTCTT CGGCAGTCAT C
```

FIG.7A.

```
   1  CGCCACTTCA ATTTTGGATT GTTGAAATTC AACTAACCAA AAAGTGCGGT
  51  TAAAATCTGT GGAGAAAATA GGTTGTAGTG AAGAACGAGG TAATTGTTCA
 101  AAAGGATAAA GCTCTCTTAA TTGGGCATTG GTTGGCGTTT CTTTTTCGGT
 151  TAATAGTAAA TTATATTCTG GACGACTATG CAATCCACCA ACAACTTTAC
 201  CGTTGGTTTT AAGCGTTAAT GTAAGTTCTT GCTCTTCTTG GCGAATACGT
 251  AATCCCATTT TTTGTTTAGC AAGAAAATGA TCGGGATAAT CATAATAGGT
 301  GTTGCCCAAA AATAAATTTT GATGTTCTAA AATCATAAAT TTTGCAAGAT
 351  ATTGTGGCAA TTCAATACCT ATTTGTGGCG AAATCGCCAA TTTTAATTCA
 401  ATTTCTTGTA GCATAATATT TCCCACTCAA ATCAACTGGT TAAATATACA
 451  AGATAATAAA AATAAATCAA GATTTTTGTG ATGACAAACA ACAATTACAA
 501  CACCTTTTTT GCAGTCTATA TGCAAATATT TTAAAAAAAT AGTATAAATC
 551  CGCCATATAA AATGGTATAA TCTTTCATCT TTCATCTTTC ATCTTTCATC
 601  TTTCATCTTT CATCTTTCAT CTTTCATCTT TCATCTTTCA TCTTTCATCT
 651  TTCATCTTTC ATCTTTCATC TTTCATCTTT CACATGAAAT GATGAACCGA
 701  GGGAAGGGAG GGAGGGGCAA GAATGAAGAG GGAGCTGAAC GAACGCAAAT
 751  GATAAAGTAA TTTAATTGTT CAACTAACCT TAGGAGAAAA TATGAACAAG
```

FIG.7B.

```
 801  ATATATCGTC TCAAATTCAG CAAACGCCTG AATGCTTTGG TTGCTGTGTC
 851  TGAATTGGCA CGGGGTTGTG ACCATTCCAC AGAAAAAGGC AGCGAAAAAC
 901  CTGCTCGCAT GAAAGTGCGT CACTTAGCGT TAAAGCCACT TTCCGCTATG
 951  TTACTATCTT TAGGTGTAAC ATCTATTCCA CAATCTGTTT TAGCAAGCGG
1001  CAATTTAACA TCGACCAAAA TGAAATGGTG CAGTTTTTAC AAGAAAACAA
1051  GTAATAAAAC CATTATCCGC AACAGTGTTG ACGCTATCAT TAATTGGAAA
1101  CAATTTAACA TCGACCAAAA TGAAATGGTG CAGTTTTTAC AAGAAAACAA
1151  CAACTCCGCC GTATTCAACC GTGTTACATC TAACCAAATC TCCCAATTAA
1201  AAGGGATTTT AGATTCTAAC GGACAAGTCT TTTTAATCAA CCCAAATGGT
1251  ATCACAATAG GTAAAGACGC AATTATTAAC ACTAATGGCT TTACGGCTTC
1301  TACGCTAGAC ATTTCTAACG AAAACATCAA GGCGCGTAAT TTCACCTTCG
1351  AGCAAACCAA AGATAAAGCG CTCGCTGAAA TTGTGAATCA CGGTTTAATT
1401  ACTGTCGGTA AAGACGGCAG TGTAAATCTT ATTGGTGGCA AAGTGAAAAA
1451  CGAGGGTGTG ATTAGCGTAA ATGGTGGCAG CATTTCTTTA CTCGCAGGGC
1501  AAAAAATCAC CATCAGCGAT ATAATAAACC CAACCATTAC TTACAGCATT
1551  GCCGCGCCTG AAAATGAAGC GGTCAATCTG GGCGATATTT TTGCCAAAGG
```

FIG.7C.

```
1601  CGGTAACATT AATGTCCGTG CTGCCACTAT TCGAAACCAA GGTAAACTTT
1651  CTGCTGATTC TGTAAGCAAA GATAAAAGCG GCAATATTGT TCTTTCCGCC
1701  AAAGAGGGTG AAGCGGAAAT TGGCGGTGTA ATTTCCGCTC AAAATCAGCA
1751  AGCTAAAGGC GGCAAGCTGA TGATTACAGG CGATAAAGTC ACATTAAAAA
1801  CAGGTGCAGT TATCGACCTT TCAGGTAAAG AAGGGGGAGA AACTTACCTT
1851  GGCGGTGACG AGCGCGGCGA AGGTAAAAAC GGCATTCAAT TAGCAAAGAA
1901  AACCTCTTTA GAAAAAGGCT CAACCATCAA TGTATCAGGC AAAGAAAAAG
1951  GCGGACGCGC TATTGTGTGG GGCGATATTG CGTTAATTGA CGGCAATATT
2001  AACGCTCAAG GTAGTGGTGA TATCGCTAAA ACCGGTGGTT TTGTGGAGAC
2051  ATCGGGGCAT TATTTATCCA TTGACAGCAA TGCAATTGTT AAAACAAAAG
2101  AGTGGTTGCT AGACCCTGAT GATGTAACAA TTGAAGCCGA AGACCCCCTT
2151  CGCAATAATA CCGGTATAAA TGATGAATTC CCAACAGGCA CCGGTGAAGC
2201  AAGCGACCCT AAAAAAAATA GCGAACTCAA AACAACGCTA ACCAATACAA
2251  CTATTTCAAA TTATCTGAAA AACGCCTGGA CAATGAATAT AACGGCATCA
2301  AGAAAACTTA CCGTTAATAG CTCAATCAAC ATCGGAAGCA ACTCCCACTT
2351  AATTCTCCAT AGTAAAGGTC AGCGTGGCGG AGGCGTTCAG ATTGATGGAG
2401  ATATTACTTC TAAAGGCGGA AATTTAACCA TTTATTCTGG CGGATGGGTT
```

FIG.7D.

2451 GATGTTCATA AAAATATTAC GCTTGATCAG GGTTTTTTAA ATATTACCGC
2501 CGCTTCCGTA GCTTTTGAAG GTGGAAATAA CAAAGCACGC GACGCGGCAA
2551 ATGCTAAAAT TGTCGCCCAG GGCACTGTAA CCATTACAGG AGAGGGAAAA
2601 GATTTCAGGG CTAACAACGT ATCTTTAAAC GGAACGGGTA AAGGTCTGAA
2651 TATCATTTCA TCAGTGAATA ATTTAACCCA CAATCTTAGT GGCACAATTA
2701 ACATATCTGG GAATATAACA ATTAACCAAA CTACGAGAAA GAACACCTCG
2751 TATTGGCAAA CCAGCCATGA TTCGCACTGG AACGTCAGTG CTCTTAATCT
2801 AGAGACAGGC GCAAATTTTA CCTTTATTAA ATACATTTCA AGCAATAGCA
2851 AAGGCTTAAC AACACAGTAT AGAAGCTCTG CAGGGGTGAA TTTTAACGGC
2901 GTAAATGGCA ACATGTCATT CAATCTCAAA GAAGGAGCGA AAGTTAATTT
2951 CAAATTAAAA CCAAACGAGA ACATGAACAC AAGCAAACCT TTACCAATTC
3001 GGTTTTTAGC CAATATCACA GCCACTGGTG GGGGCTCTGT TTTTTTTGAT
3051 ATATATGCCA ACCATTCTGG CAGAGGGGCT GAGTTAAAAA TGAGTGAAAT
3101 TAATATCTCT AACGGCGCTA ATTTTACCTT AAATTCCCAT GTTCGCGGCG
3151 ATGACGCTTT TAAAATCAAC AAAGACTTAA CCATAAATGC AACCAATTCA
3201 AATTTCAGCC TCAGACAGAC GAAAGATGAT TTTTATGACG GGTACGCACG

FIG.7E.

```
3251  CAATGCCATC AATTCAACCT ACAACATATC CATTCTGGGC GGTAATGTCA
3301  CCCTTGGTGG ACAAAACTCA AGCAGCAGCA TTACGGGGAA TATTACTATC
3351  GAGAAAGCAG CAAATGTTAC GCTAGAAGCC AATAACGCCC CTAATCAGCA
3401  AAACATAAGG GATAGAGTTA TAAAACTTGG CAGCTTGCTC GTTAATGGGA
3451  GTTTAAGTTT AACTGGCGAA AATGCAGATA TTAAAGGCAA TCTCACTATT
3501  TCAGAAAGCG CCACTTTTAA AGGAAAGACT AGAGATACCC TAAATATCAC
3551  CGGCAATTTT ACCAATAATG GCACTGCCGA AATTAATATA ACACAAGGAG
3601  TGGTAAAACT TGGCAATGTT ACCAATGATG GTGATTTAAA CATTACCACT
3651  CACGCTAAAC GCAACCAAAG AAGCATCATC GGCGGAGATA TAATCAACAA
3701  AAAAGGAAGC TTAAATATTA CAGACAGTAA TAATGATGCT GAAATCCAAA
3751  TTGGCGGCAA TATCTCGCAA AAAGAAGGCA ACCTCACGAT TTCTTCCGAT
3801  AAAATTAATA TCACCAAACA GATAACAATC AAAAAGGGTA TTGATGGAGA
3851  GGACTCTAGT TCAGATGCGA CAAGTAATGC CAACCTAACT ATTAAAACCA
3901  AAGAATTGAA ATTGACAGAA GACCTAAGTA TTTCAGGTTT CAATAAAGCA
3951  GAGATTACAG CCAAAGATGG TAGAGATTTA ACTATTGGCA ACAGTAATGA
4001  CGGTAACAGC GGTGCCGAAG CCAAAACAGT AACTTTTAAC AATGTTAAAG
```

FIG.7F.

```
4051  ATTCAAAAAT CTCTGCTGAC GGTCACAATG TGACACTAAA TAGCAAAGTG
4101  AAAACATCTA GCAGCAATGG CGGACGTGAA AGCAATAGCG ACAACGATAC
4151  CGGCTTAACT ATTACTGCAA AAAATGTAGA AGTAAACAAA GATATTACTT
4201  CTCTCAAAAC AGTAAATATC ACCGCGTCGG AAAAGGTTAC CACCACAGCA
4251  GGCTCGACCA TTAACGCAAC AAATGGCAAA GCAAGTATTA CAACCAAAAC
4301  AGGTGATATC AGCGGTACGA TTTCCGGTAA CACGGTAAGT GTTAGCGCGA
4351  CTGGTGATTT AACCACTAAA TCCGGCTCAA AAATTGAAGC GAAATCGGGT
4401  GAGGCTAATG TAACAAGTGC AACAGGTACA ATTGGCGGTA CAATTTCCGG
4451  TAATACGGTA AATGTTACGG CAAACGCTGG CGATTTAACA GTTGGGAATG
4501  GCGCAGAAAT TAATGCGACA GAAGGAGCTG CAACCTTAAC CGCAACAGGG
4551  AATACCTTGA CTACTGAAGC CGGTTCTAGC ATCACTTCAA CTAAGGGTCA
4601  GGTAGACCTC TTGGCTCAGA ATGGTAGCAT CGCAGGAAGC ATTAATGCTG
4651  CTAATGTGAC ATTAAATACT ACAGGCACCT TAACCACCGT GGCAGGCTCG
4701  GATATTAAAG CAACCAGCGG CACCTTGGTT ATTAACGCAA AAGATGCTAA
4751  GCTAAATGGT GATGCATCAG GTGATAGTAC AGAAGTGAAT GCAGTCAACG
4801  ACTGGGGATT TGGTAGTGTG ACTGCGGCAA CCTCAAGCAG TGTGAATATC
4851  ACTGGGGATT TAAACACAGT AAATGGGTTA AATATCATTT CGAAAGATGG
```

FIG.7G.

```
4901  TAGAAACACT GTGCGCTTAA GAGGCAAGGA AATTGAGGTG AAATATATCC
4951  AGCCAGGTGT AGCAAGTGTA GAAGAAGTAA TTGAAGCGAA ACGCGTCCTT
5001  GAAAAAGTAA AAGATTTATC TGATGAAGAA AGAGAAACAT TAGCTAAACT
5051  TGGTGTAAGT GCTGTACGTT TTGTTGAGCC AAATAATACA ATTACAGTCA
5101  ATACACAAAA TGAATTTACA ACCAGACCGT CAAGTCAAGT GATAATTTCT
5151  GAAGGTAAGG CGTGTTTCTC AAGTGGTAAT GGCGCACGAG TATGTACCAA
5201  TGTTGCTGAC GATGGACAGC CGTAGTCAGT AATTGACAAG GTAGATTTCA
5251  TCCTGCAATG AAGTCATTTT ATTTTCGTAT TATTTACTGT GTGGGTTAAA
5301  GTTCAGTACG GGCTTTACCC ATCTTGTAAA AAATTACGGA GAATACAATA
5351  AAGTATTTTT AACAGGTTAT TATTATGAAA AATATAAAAA GCAGATTAAA
5401  ACTCAGTGCA ATATCAGTAT TGCTTGGCCT GGCTTCTTCA TCATTGTATG
5451  CAGAAGAAGC GTTTTTAGTA AAAGGCTTTC AGTTATCTGG TGCACTTGAA
5501  ACTTTAAGTG AAGACGCCCA ACTGTCTGTA GCAAAATCTT TATCTAAATA
5551  CCAAGGCTCG CAAACTTTAA CAAACCTAAA AACAGCACAG CTTGAATTAC
5601  AGGCTGTGCT AGATAAGATT GAGCCAAATA AATTTGATGT GATATTGCCG
5651  CAACAAACCA TTACGGATGG CAATATCATG TTTGAGCTAG TCTCGAAATC
```

FIG.7H.

```
5701  AGCCGCAGAA AGCCAAGTTT TTTATAAGGC GAGCCAGGGT TATAGTGAAG
5751  AAAATATCGC TCGTAGCCTG CCATCTTTGA AACAAGGAAA AGTGTATGAA
5801  GATGGTCGTC AGTGGTTCGA TTTGCGTGAA TTTAATATGG CAAAAGAAAA
5851  CCCGCTTAAG GTTACCCGTG TACATTACGA ACTAAACCCT AAAAACAAAA
5901  CCTCTAATTT GATAATTGCG GGCTTCTCGC CTTTTGGTAA AACGCGTAGC
5951  TTTATTTCTT ATGATAATTT CGGCGCGAGA GAGTTTAACT ACCAACGTGT
6001  AAGCTTGGGT TTTGTTAATG CCAATTTAAC TGGTCATGAT GATGTGTTAA
6151  TTATACCAGT ATGAGTTATG CTGATTCTAA TGATATCGAC GGCTTACCAA
6201  GTGCGATTAA TCGTAAATTA TCAAAAGGTC AATCTATCTC TGCGAATCTG
6251  AAATGGAGTT ATTATCTCCC AACATTTAAC CTTGGCATGG AAGACCAATT
6301  TAAAATTAAT TTAGGCTACA ACTACCGCCA TATTAATCAA ACCTCCGCGT
6351  TAAATCGCTT GGGTGAAACG AAGAAAAAAT TTGCAGTATC AGGCGTAAGT
6401  GCAGGCATTG ATGGACATAT CCAATTTACC CCTAAAACAA TCTTTAATAT
6451  TGATTTAACT CATCATTATT ACGCGAGTAA ATTACCAGGC TCTTTTGGAA
6501  TGGAGCGCAT TGGCGAAACA TTTAATCGCA GCTATCACAT TAGCACAGCC
6551  AGTTTAGGGT TGAGTCAAGA GTTTGCTCAA GGTTGGCATT TTAGCAGTCA
6601  ATTATCAGGT CAATTTACTC TACAAGATAT TAGCAGTATA GATTTATTCT
```

FIG.7I.

```
6651  CTGTAACAGG TACTTATGGC GTCAGAGGCT TTAAATACGG CGGTGCAAGT
6701  GGTGAGCGCG GTCTTGTATG GCGTAATGAA TTAAGTATGC CAAAATACAC
6751  CCGCTTCCAA ATCAGCCCTT ATGCGTTTTA TGATGCAGGT CAGTTCCGTT
6801  ATAATAGCGA AAATGCTAAA ACTTACGGCG AAGATATGCA CACGGTATCC
6851  TCTGCGGGTT TAGGCATTAA AACCTCTCCT ACACAAAACT TAAGCCTAGA
6901  TGCTTTTGTT GCTCGTCGCT TTGCAAATGC CAATAGTGAC AATTTGAATG
6951  GCAACAAAAA ACGCACAAGC TCACCTACAA CCTTCTGGGG GAGATTAACA
7001  TTCAGTTTCT AACCCTGAAA TTTAATCAAC TGGTAAGCGT TCCGCCTACC
7051  AGTTTATAAC TATATGCTTT ACCCGCCAAT TTACAGTCTA TAGGCAACCC
7101  TGTTTTTACC CTTATATATC AAATAAACAA GCTAAGCTGA GCTAAGCAAA
7151  CCAAGCAAAC TCAAGCAAGC CAAGTAATAC TAAAAAAACA ATTTATATGA
7201  TAAACTAAAG TATACTCCAT GCCATGGCGA TACAAGGGAT TTAATAATAT
7251  GACAAAAGAA AATTTGCAAA ACGCTCCTCA AGATGCGACC GCTTTACTTG
7301  CGGAATTAAG CAACAATCAA ACTCCCCTGC GAATATTTAA ACAACCACGC
7351  AAGCCCAGCC TATTACGCTT GGAACAACAT ATCGCAAAAA AAGATTATGA
7401  GTTTGCTTGT CGTGAATTAA TGGTGATTCT GGAAAAAATG GACGCTAATT
```

FIG.7J.

```
7451  TTGGAGGCGT TCACGATATT GAATTTGACG CACCCGCTCA GCTGGCATAT
7501  CTACCCGAAA AATTACTAAT TTATTTTGCC ACTCGTCTCG CTAATGCAAT
7551  TACAACACTC TTTTCCGACC CCGAATTGGC AATTTCTGAA GAAGGGGCGT
7601  TAAAGATGAT TAGCCTGCAA CGCTGGTTGA CGCTGATTTT TGCCTCTTCC
7651  CCCTACGTTA ACGCAGACCA TATTCTCAAT AAATATAATA TCAACCCAGA
7701  TTCCGAAGGT GGCTTTCATT TAGCAACAGA CAACTCTTCT ATTGCTAAAT
7751  TCTGTATTTT TTACTTACCC GAATCCAATG TCAATATGAG TTTAGATGCG
7801  TTATGGGCAG GGAATCAACA ACTTTGTGCT TCATTGTGTT TTGCGTTGCA
7851  GTCTTCACGT TTTATTGGTA CCGCATCTGC GTTTCATAAA AGAGCGGTGG
7901  TTTTACAGTG GTTTCCTAAA AAACTCGCCG AAATTGCTAA TTTAGATGAA
7951  TTGCCTGCAA ATATCCTTCA TGATGTATAT ATGCACTGCA GTTATGATTT
8001  AGCAAAAAAC AAGCACGATG TTAAGCGTCC ATTAAACGAA CTTGTCCGCA
8051  AGCATATCCT CACGCAAGGA TGGCAAGACC GCTACCTTTA CACCTTAGGT
8101  AAAAAGGACG GCAAACCTGT GATGATGGTA CTGCTTGAAC ATTTTAATTC
8151  GGGACATTCG ATTTATCGTA CACATTCAAC TTCAATGATT GCTGCTCGAG
8201  AAAAATTCTA TTTAGTCGGC TTAGGCCATG AGGGCGTTGA TAAAATAGGT
```

FIG.7K.

```
8251  CGAGAAGTGT TTGACGAGTT CTTTGAAATC AGTAGCAATA ATATAATGGA
8301  GAGACTGTTT TTTATCCGTA AACAGTGCGA AACTTTCCAA CCCGCAGTGT
8351  TCTATATGCC AAGCATTGGC ATGGATATTA CCACGATTTT TGTGAGCAAC
8401  ACTCGGCTTG CCCCTATTCA AGCTGTAGCC CTGGGTCATC CTGCCACTAC
8451  GCATTCTGAA TTTATTGATT ATGTCATCGT AGAAGATGAT TATGTGGGCA
8501  GTGAAGATTG TTTCAGCGAA ACCCTTTTAC GCTTACCCAA AGATGCCCTA
8551  CCTTATGTAC CTTCTGCACT CGCCCCACAA AAAGTGGATT ATGTACTCAG
8601  GGAAAACCCT GAAGTAGTCA ATATCGGTAT TGCCGCTACC ACAATGAAAT
8651  TAAACCCTGA ATTTTTGCTA ACATTGCAAG AAATCAGAGA TAAAGCTAAA
8701  GTCAAAATAC ATTTTCATTT CGCACTTGGA CAATCAACAG GCTTGACACA
8751  CCCTTATGTC AAATGGTTTA TCGAAAGCTA TTTAGGTGAC GATGCCACTG
8801  CACATCCCCA CGCACCTTAT CACGATTATC TGGCAATATT GCGTGATTGC
8851  GATATGCTAC TAAATCCGTT TCCTTTCGGT AATACTAACG GCATAATTGA
8901  TATGGTTACA TTAGGTTTAG TTGGTGTATG CAAAACGGGG GATGAAGTAC
8951  ATGAACATAT TGATGAAGGT CTGTTTAAAC GCTTAGGACT ACCAGAATGG
9001  CTGATAGCCG ACACACGAGA AACATATATT GAATGTGCTT TGCGTCTAGC
9051  AGAAAACCAT CAAGAACGCC TTGAACTCCG TCGTTACATC ATAGAAAACA
```

FIG.7L.

```
9101  ACGGCTTACA AAAGCTTTTT ACAGGCGACC CTCGTCCATT GGGCAAAATA
9151  CTGCTTAAGA AAACAAATGA ATGGAAGCGG AAGCACTTGA GTAAAAAATA
9201  ACGGTTTTTT AAAGTAAAAG TGCGGTTAAT TTTCAAAGCG TTTTAAAAAC
9251  CTCTCAAAAA TCAACCGCAC TTTTATCTTT ATAACGATCC CGCACGCTGA
9301  CAGTTTATCA GCCTCCCGCC ATAAAACTCC GCCTTTCATG GCGGAGATTT
9351  TAGCCAAAAC TGGCAGAAAT TAAAGGCTAA AATCACCAAA TTGCACCACA
9401  AAATCACCAA TACCCACAAA AAA
```

FIG.8A

FIG 8B

801 TGCCAAAGGT GGTAACATTA ATGTCCGCGC TGCCACTATT CGCAATAAAG
GTAAACTTTC TGCCGACTCT GTAAGCAAAG ATAAAAGTGG TAACATTGTT

901 CTCTCTGCCA AAGAAGGTGA AGCGGAAATT GGCGGTGTAA TTTCCGCTCA
AAATCAGCAA GCCAAAGGTG GTAAGTTGAT GATTACAGGC GATAAAGTTA

1001 CATTGAAAAC GGGTGCAGTT ATCGACCTTT CGGGTAAAGA AGGGGGAGAA
ACTTATCTTG GCGGTGACGA GCGTGGCGAA GGTAAAAACG GCATTCAATT

1101 AGCAAAGAAA ACCACTTTAG AAAAAGGCTC AACAATTAAT GTGTCAGGTA
AAGAAAAAGG TGGGCGCGCT ATTGTATGGG GCGATATTGC GTTAATTGAC

1201 GGCAATATTA ATGCCCAAGG TAAAGATATC GCTAAAACTG GTGGTTTTGT
GGAGACGTCG GGGCATTACT TATCCATTGA TGATAACGCA ATTGTTAAAA

1301 CAAAAGAATG GCTACTAGAC CCAGAGAATG TGACTATTGA AGCTCCTTCC
GCTTCTCGCG TCGAGCTGGG TGCCGATAGG AATTCCCACT CGGCAGAGGT

1401 GATAAAAGTG ACCCTAAAAA AAAATAACAC CTCCTTGACA ACACTAACCA
ATACAACCAT TTCAAATCTT CTGAAAAGTG CCCACGTGGT GAACATAACG

1501 GCAAGGAGAA AACTTACCGT TAATAGCTCT ATCAGTATAG AAAGAGGCTC
CCACTTAATT CTCCACAGTG AAGGTCAGGG CGGTCAAGGT GTTCAGATTG

FIG.8C

1601 ATAAAGATAT TACTTCTGAA GGCGGAAATT TAACCATTTA TTCTGGCGGA
TGGGTTGATG TTCATAAAAA TATTACGCTT GGTAGCGGCT TTTTAAACAT

1701 CACAACTAAA GAAGGAGATA TCGCCTTCGA AGACAAGTCT GGACGGAACA
ACCTAACCAT TACAGCCCAA GGGACCATCA CCTCAGGTAA TAGTAACGGC

1801 TTTAGATTTA ACAACGTCTC TCTAAACAGC CTTGGCGGAA AGCTGAGCTT
TACTGACAGC AGAGAGGACA GAGGTAGAAG AACTAAGGGT AATATCTCAA

1901 ACAAATTTGA CGGAACGTTA AACATTTCCG GAACTGTAGA TATCTCAATG
AAAGCACCCA AGTCAGCTG GTTTTACAGA GACAAAGGAC GCACCTACTG

2001 GAACGTAACC ACTTTAAATG TTACCTCGGG TAGTAAATTT AACCTCTCCA
TTGACAGCAC AGGAAGTGGC TCAACAGGTC CAAGCATACG CAATGCAGAA

2101 TTAAATGGCA TAACATTTAA TAAAGCCACT TTTAATATCG CACAAGGCTC
AACAGCTAAC TTTAGCATCA AGGCATCAAT AATGCCCTTT AAGAGTAACG

2201 CTAACTACGC ATTATTTAAT GAAGATATTT CAGTCTCAGG GGGGGGTAGC
CTTAATTTCA AACTTAACGC CTCATCTAGC AACATACAAA CCCCTGGCGT

2301 AATTATAAAA TCTCAAAACT TTAATGTCTC AGGAGGGTCA ACTTTAAATC
TCAAGGCTGA AGGTTCAACA GAAACCGCTT TTTCAATAGA AAATGATTTA

FIG.8D

2401 AACTTAAACG CCACCGGTGG CAATATAACA ATCAGACAAG TCGAGGGTAC
CGATTCACGC GTCAACAAAG GTGTCGCAGC CAAAAAAAAC ATAACTTTTA

2501 AAGGGGGTAA TATCACCTTC GGCTCTCAAA AAGCCACAAC AGAAATCAAA
GGCAATGTTA CCATCAATAA AAACACTAAC GCTACTCTTT GTGGTGCGAA

2601 TTTTGCCGAA AACAAATCGC CTTTAAATAT AGCAGGAAAT GTTATTAATA
ATGGCAACCT TACCACTGCC GGCTCCATTA TCAATATAGC CGGAAATCTT

2701 ACTGTTTCAA AAGGCGCTAA CCTTCAAGCT ATAACAAATT ACACTTTTAA
TGTAGCCGGC TCATTTGACA ACAATGGCGC TTCAAACATT TCCATTGCCA

2801 GAGGAGGGGC TAAATTTAAA GATATCAATA ACACCAGTAG CTTAAATATT
ACCACCAACT CTGATACCAC TTACCGCACC ATTATAAAAG GCAATATATC

2901 CAACAAATCA GGTGATTTGA ATATTATTGA TAAAAAAAGC GACGCTGAAA
TCCAAATTGG CGGCAATATC TCACAAAAAG AAGGCAATCT CACAATTTCT

3001 TCTGATAAAG TAAATATTAC CAATCAGATA ACAATCAAAG CAGGCGTTGA
AGGGGGGCGT TCTGATTCAA GTGAGGCAGA AAATGCTAAC CTAACTATTC

3101 AAACCAAAGA GTTAAAATTG GCAGGAGACC TAAATATTTC AGGCTTTAAT
AAAGCAGAAA TTACAGCTAA AAATGGCAGT GATTTAACTA TTGGCAATGC

FIG.8E

3201 TAGCGGTGGT AATGCTGATG CTAAAAAAGT GACTTTTGAC AAGGTTAAAG

ATTCAAAAAT CTCGACTGAC GGTCACAATG TAACACTAAA TAGCGAAGTG

3301 AAAACGTCTA ATGGTAGTAG CAATGCTGGT AATGATAACA GCACCGGTTT

AACCATTTCC GCAAAAGATG TAACGGTAAA CAATAACGTT ACCTCCCACA

3401 AGACAATAAA TATCTCTGCC GCAGCAGGAA ATGTAACAAC CAAAGAAGGC

ACAACTATCA ATGCAACCAC AGGCAGCGTG GAAGTAACTG CTCAAAATGG

3501 TACAATTAAA GGCAACATTA CCTCGCAAAA TGTAACAGTG ACAGCAACAG

AAAATCTTGT TACCACAGAG AATGCTGTCA TTAATGCAAC CAGCGGCACA

3601 GTAAACATTA GTACAAAAAC AGGGGATATT AAAGGTGGAA TTGAATCAAC

TTCCGGTAAT GTAAATATTA CAGCGAGCGG CAATACACTT AAGGTAAGTA

3701 ATATCACTGG TCAAGATGTA ACAGTAACAG CGGATGCAGG AGCCTTGACA

ACTACAGCAG GCTCAACCAT TAGTGCGACA ACAGGCAATG CAAATATTAC

3801 AACCAAAACA GGTGATATCA ACGGTAAAGT TGAATCCAGC TCCGGCTCTG

TAACACTTGT TGCAACTGGA GCAACTCTTG CTGTAGGTAA TATTTCAGGT

3901 AACACTGTTA CTATTACTGC GGATAGCGGT AAATTAACCT CCACAGTAGG

TTCTACAATT AATGGGACTA ATAGTGTAAC CACCTCAAGC CAATCAGGCG

FIG.8F

4001 ATATTGAAGG TACAATTTCT GGTAATACAG TAAATGTTAC AGCAAGCACT
GGTGATTTAA CTATTGGAAA TAGTGCAAAA GTTGAAGCGA AAAATGGAGC

4101 TGCAACCTTA ACTGCTGAAT CAGGCAAATT AACCACCCAA ACAGGCTCTA
GCATTACCTC AAGCAATGGT CAGACAACTC TTACAGCCAA GGATAGCAGT

!201 ATCGCAGGAA ACATTAATGC TGCTAATGTG ACGTTAAATA CCACAGGCAC
TTTAACTACT ACAGGGGATT CAAAGATTAA CGCAACCAGT GGTACCTTAA

.301 CAATCAATGC AAAAGATGCC AAATTAGATG GTGCTGCATC AGGTGACCGC
ACAGTAGTAA ATGCAACTAA CGCAAGTGGC TCTGGTAACG TGACTGCGAA

401 AACCTCAAGC AGCGTGAATA TCACCGGGGA TTTAAACACA ATAAATGGGT
TAAATATCAT TTCGGAAAAT GGTAGAAACA CTGTGCGCTT AAGAGGCAAG

501 GAAATTGATG TGAAATATAT CCAACCAGGT GTAGCAAGCG TAGAAGAGGT
AATTGAAGCG AAACGCGTCC TTGAGAAGGT AAAAGATTTA TCTGATGAAG

501 AAAGAGAAAC ACTAGCCAAA CTTGGTGTAA GTGCTGTACG TTTCGTTGAG
CCAAATAATG CCATTACGGT TAATACACAA AACGAGTTTA CAACCAAACC

'01 ATCAAGTCAA GTGACAATTT CTGAAGGTAA GGCGTGTTTC TCAAGTGGTA
ATGGCGCACG AGTATGTACC AATGTTGCTG ACGATGGACA GCAG

FIG.9A

```
1 ATGAACAAGA TATATCGTCT CAAATTCAGC AAACGCCTGA ATGCTTTGGT
                                    TGCTGTGTCT GAATTGACAC GGGGTTGTGA CCATTCCACA GAAAAAGGCA

101 GTGAAAAACC TGTTCGTACG AAAGTACGCC ACTTGGCGTT AAAGCCACTT
                                    TCCGCTATAT TGCTATCTTT GGGCATGGCA TCCATTCCGC AATCTGTTTT

201 AGCGAGCGGT TTACAGGGAA TGAGCGTCGT ACACGGTACA GCAACCATGC
                                    AAGTAGACGG CAATAAAACC ACTATCCGTA ATAGCGTCAA TGCTATCATC

301 AATTGGAAAC AATTTAACAT TGACCAAAAT GAAATGGTGC AGTTTTTACA
                                    AGAAAGCAGC AACTCTGCCG TTTTCAACCG TGTTACATCT GACCAAATCT

401 CCCAATTAAA AGGGATTTTA GATTCTAACG GACAAGTCTT TTTAATCAAC
                                    CCAAATGGTA TCACAATAGG TAAAGACGCA ATTATTAACA CTAATGGCTT

501 TACTGCTTCT ACGCTAGACA TTTCTAACGA AAACATCAAG GCGCGTAATT
                                    TCACCCTTGA GCAAACCAAG GATAAAGCAC TCGCTGAAAT CGTGAATCAC

601 GGTTTAATTA CCGTTGGTAA AGACGGTAGC GTAAACCTTA TTGGTGGCAA
                                    AGTGAAAAAC GAGGGCGTGA TTAGCGTAAA TGGCGGTAGT ATTTCTTTAC

701 TTGCAGGGCA AAAAATCACC ATCAGCGATA TAATAAATCC AACCATCACT
                                    TACAGCATTG CTGCACCTGA AAACGAAGCG ATCAATCTGG GCGATATTTT
```

FIG.9B

801 TGCCAAAGGT GGTAACATTA ATGTCCGCGC TGCCACTATT CGCAATAAAG
GTAAACTTTC TGCCGACTCT GTAAGCAAAG ATAAAAGTGG TAACATTGTT

901 CTCTCTGCCA AAGAAGGTGA AGCGGAAATT GGCGGTGTAA TTTCCGCTCA
AAATCAGCAA GCCAAAGGTG GTAAGTTGAT GATTACAGGT GATAAAGTCA

1001 CATTAAAAAC AGGTGCAGTT ATCGACCTTT CAGGTAAAGA AGGGGGAGAG
ACTTATCTTG GCGGTGATGA GCGTGGCGAA GGTAAAAATG GTATTCAATT

1101 AGCGAAGAAA ACCTCTTTAG AAAAAGGCTC GACAATTAAT GTATCAGGCA
AAGAAAAAGG CGGGCGCGCT ATTGTATGGG GCGATATTGC ATTAATTAAT

1201 GGTAACATTA ATGCTCAAGG TAGCGATATT GCTAAAACTG GCGGCTTTGT
GGAAACATCA GGACATGACT TATCCATTGG TGATGATGTG ATTGTTGACG

1301 CTAAAGAGTG GTTATTAGAC CCAGATGATG TGTCCATTGA AACTCTTACA
TCTGGACGCA ATAATACCGG CGAAAACCAA GGATATACAA CAGGAGATGG

1401 GACTAAAGAG TCACCTAAAG GTAATAGTAT TTCTAAACCT ACATTAACAA
ACTCAACTCT TGAGCAAATC CTAAGAAGAG GTTCTTATGT TAATATCACT

1501 GCTAATAATA GAATTTATGT TAATAGCTCC ATCAACTTAT CTAATGGCAG
TTTAACACTT CACACTAAAC GAGATGGAGT TAAAATTAAC GGTGATATTA

FIG.9C

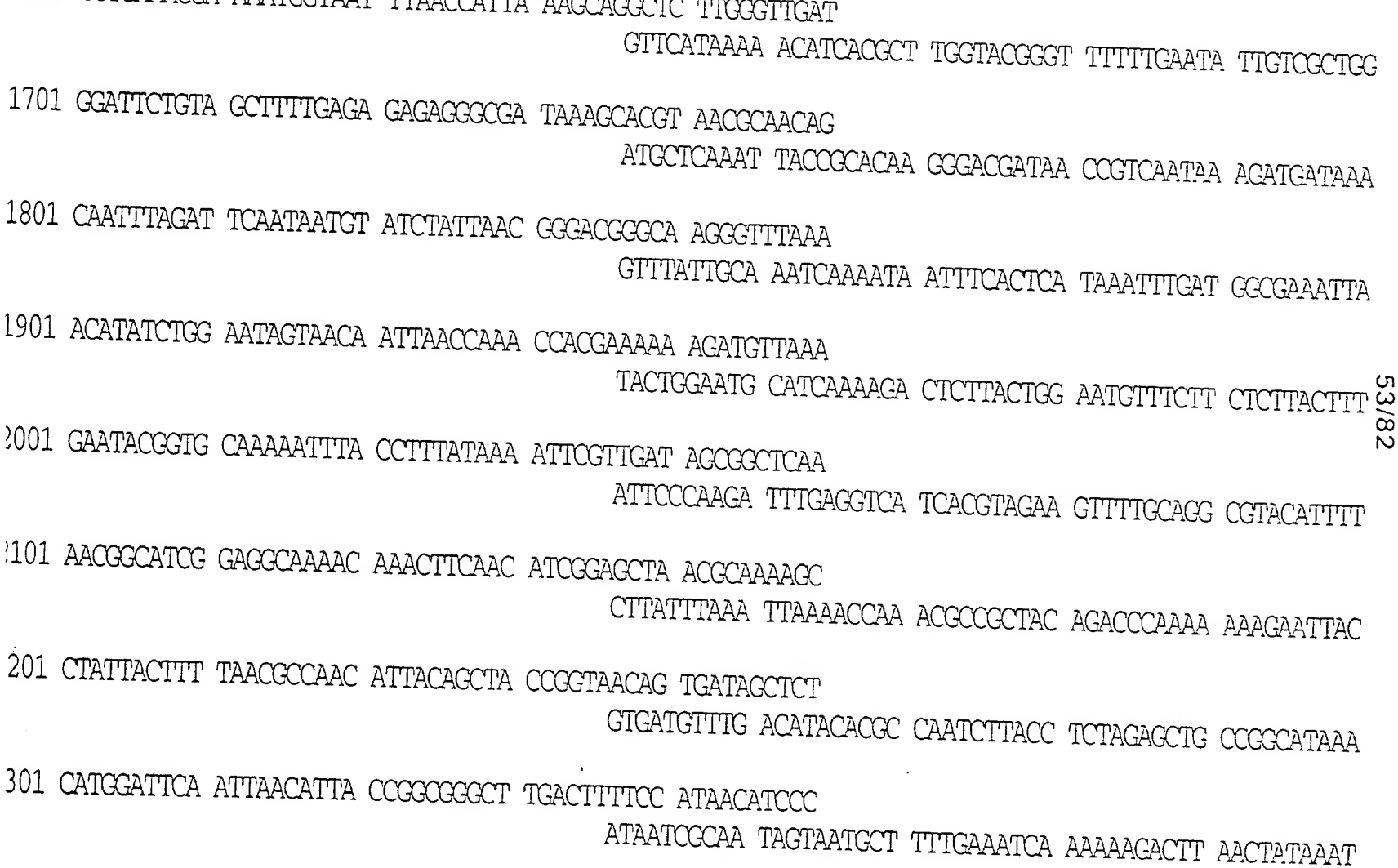


FIG.9D

2401 GCAACTGGCT CGAATTTTAG TCTTAAGCAA ACGAAAGATT CTTTTTATAA
TGAATACAGC AAACACGCCA TTAACTCAAG TCATAATCTA ACCATTCTTG
2501 GCGGCAATGT CACTCTAGGT GGGGAAAATT CAAGCAGTAG CATTACGGGC
AATATCAATA TCACCAATAA AGCAAATGTT ACATTACAAG CTGACACCAG
2601 CAACAGCAAC ACAGGCTTGA AGAAAAGAAC TCTAACTCTT GGCAATATAT
CTGTTGAGGG GAATTTAAGC CTAACTGGTG CAAATGCAAA CATTGTCGGC
2701 AATCTTTCTA TTGCAGAAGA TTCCACATTT AAAGGAGAAG CCAGTGACAA
CCTAAACATC ACCGGCACCT TTACCAACAA CGGTACCGCC AACATTAATA
2801 TAAAACAAGG AGTGGTAAAA CTCCAAGGCG ATATTATCAA TAAAGGTGGT
TTAAATATCA CTACTAACGC CTCAGGCACT CAAAAAACCA TTATTAACGG
2901 AAATATAACT AACGAAAAAG GCGACTTAAA CATCAAGAAT ATTAAAGCCG
ACGCCGAAAT CCAAATTGGC GGCAATATCT CACAAAAAGA AGGCAATCTC
3001 ACAATTTCTT CTGATAAAGT AAATATTACC AATCAGATAA CAATCAAAGC
AGGCGTTGAA GGGGGGCGTT CTGATTCAAG TGAGGCAGAA AATGCTAACC
3101 TAACTATTCA AACCAAAGAG TTAAAATTGG CAGGAGACCT AAATATTTCA
GGCTTTAATA AAGCAGAAAT TACAGCTAAA AATGGCAGTG ATTTAACTAT
3201 TGGCAATGCT AGCGGTGGTA ATGCTGATGC TAAAAAAGTG ACTTTTGACA
AGGTTAAAGA TTCAAAAATC TCGACTGACG GTCACAATGT AACACTAAAT

FIG.9E

3301 AGCGAAGTGA AAACGTCTAA TGGTAGTAGC AATGCTGGTA ATGATAACAG
CACCGGTTTA ACCATTTCCG CAAAAGATGT AACGGTAAAC AATAACGTTA

3401 CCTCCCACAA GACAATAAAT ATCTCTGCCG CAGCAGGAAA TGTAACAACC
AAAGAAGGCA CAACTATCAA TGCAACCACA GGCAGCGTGG AAGTAACTGC

3501 TCAAAATGGT ACAATTAAAG GCAACATTAC CTCGCAAAAT GTAACAGTGA
CAGCAACAGA AAATCTTGTT ACCACAGAGA ATGCTGTCAT TAATGCAACC

3601 AGCGGCACAG TAAACATTAG TACAAAAACA GGGGATATTA AAGGTGGAAT
TGAATCAACT TCCGGTAATG TAAATATTAC AGCGAGCGGC AATACACTTA

3701 AGGTAAGTAA TATCACTGGT CAAGATGTAA CAGTAACAGC GGATGCAGGA
GCCTTGACAA CTACAGCAGG CTCAACCATT AGTGCGACAA CAGGCAATGC

3801 AAATATTACA ACCAAAACAG GTGATATCAA CGGTAAAGTT GAATCCAGCT
CCGGCTCTGT AACACTTGTT GCAACTGGAG CAACTCTTGC TGTAGGTAAT

3901 ATTTCAGGTA ACACTGTTAC TATTACTGCG GATAGCGGTA AATTAACCTC
CACAGTAGGT TCTACAATTA ATGGGACTAA TAGTGTAACC ACCTCAAGCC

4001 AATCAGGCGA TATTGAAGGT ACAATTTCTG GTAATACAGT AAATGTTACA
GCAAGCACTG GTGATTTAAC TATTGGAAAT AGTGCAAAAG TTGAAGCGAA

FIG.9F

```
4101 AAATGGAGCT GCAACCTTAA CTGCTGAATC AGGCAAATTA ACCACCCAAA
                                    CAGGCTCTAG CATTACCTCA AGCAATGGTC AGACAACTCT TACAGCCAAG
4201 GATAGCAGTA TCGCAGGAAA CATTAATGCT GCTAATGTGA CGTTAAATAC
                                    CACAGGCACT TTAACTACTA CAGGGGATTC AAAGATTAAC GCAACCAGTG
4301 GTACCTTAAC AATCAATGCA AAAGATGCCA AATTAGATGG TGCTGCATCA
                                    GGTGACCGCA CAGTAGTAAA TGCAACTAAC GCAAGTGGCT CTGGTAACGT
4401 GACTGCGAAA ACCTCAAGCA GCGTGAATAT CACCGGGGAT TTAAACACAA
                                    TAAATGGGTT AAATATCATT TCGGAAAATG GTAGAAACAC TGTGCGCTTA
4501 AGAGGCAAGG AAATTGATGT GAAATATATC CAACCAGGTG TAGCAAGCGT
                                    AGAAGAGGTA ATTGAAGCGA AACGCGTCCT TGAGAAGGTA AAAGATTTAT
4601 CTGATGAAGA AAGAGAAACA CTAGCCAAAC TTGGTGTAAG TGCTGTACGT
                                    TTCGTTGAGC CAAATAATGC CATTACGGTT AATACACAAA ACGAGTTTAC
4701 AACCAAACCA TCAAGTCAAG TGACAATTTC TGAAGGTAAG GCGTGTTTCT
                                    CAAGTGGTAA TGGCGCACGA GTATGTACCA ATGTTGCTGA CGATGGACAG
4801 CAG
```

FIG.10A

COMPARISON OF DERIVED AMINO ACID SEQUENCE

```
                1                                                  50
Hmw3com         MNKIYRLKFS KRLNALVAVS ELTRGCDHST EKGSEKPVRT KVRHLALKPL
Hmw4com         MNKIYRLKFS KRLNALVAVS ELTRGCDHST EKGSEKPVRT KVRHLALKPL
Hmw1com         MNKIYRLKFS KRLNALVAVS ELTRGCDHST EKGSEKPVRT KVRHLALKPL
Hmw2com         MNKIYRLKFS KRLNALVAVS ELTRGCDHST EKGSEKPVRT KVRHLALKPL

               51                                                 100
Hmw3com         SAILLSLGMA SIPQSVLASG LQGMSVVHGT ATMQVDGNKT TIRNSVNAII
Hmw4com         SAILLSLGMA SIPQSVLASG LQGMSVVHGT ATMQVDGNKT TIRNSVNAII
Hmw2com         SAMLLSLGVT SIPQSVLASG LQGMSVVHGT ATMQVDGNKT TIRNSVNAII
Hmw2com         SAMLLSLGVT SIPQSVLASG LQGMSVVHGT ATMQVDGNKT TIRNSVNAII
```

FIG.10B

	101				150
Hmw3com	NWKQFNIDQN	EMEQFLQESS	NSAVFNRVTS	DQISQLKSIL	DSNSQVFLIN
Hmw4com	NWKQFNIDQN	EMEQFLQESS	NSAVFNRVTS	DQISQLKGIL	DSNSQVFLIN
Hmw1com	NWKQFNIDQN	EMVQFLQENN	NSAVFNRVTS	NQISQLKGIL	DSNSQVFLIN
Hmw2com	NWKQFNIDQN	EMVQFLQENN	NSAVFNRVTS	NQISQLKGIL	DSNSQVFLIN

	151				200
Hmw3com	PNGITIGKDA	IINTNSFTAS	TLDISNENIK	ARNFTLEQTK	DKALAEIVNH
Hmw4com	PNGITIGKDA	IINTNSFTAS	TLDISNENIK	ARNFTLEQTK	DKALAEIVNH
Hmw1com	PNGITIGKDA	IINTNSFTAS	TLDISNENIK	ARNFTLEQTK	DKALAEIVNH
Hmw2com	PNGITIGKDA	IINTNSFTAS	TLDISNENIK	ARNFTLEQTK	DKALAEIVNH

FIG.10C

	201				250
Hmw3com	GLITVGKDGS	VNLIGGKVKN	EGVISVNGGS	ISLLAGQKIT	ISDIINPTIT
Hmw4com	GLITVGKDGS	VNLIGGKVKN	EGVISVNGGS	ISLLAGQKIT	ISDIINPTIT
Hmw1com	GLITVGKDGS	VNLIGGKVKN	EGVISVNGGS	ISLLAGQKIT	ISDIINPTIT
Hmw2com	GLITVGKDGS	VNLIGGKVKN	EGVISVNGGS	ISLLAGQKIT	ISDIINPTIT

	251				300
Hmw3com	YSIAAPENEA	INLGDIFAKG	GNIINVRAATI	RNKGKLSADS	VSKDKSGNIV

FIG.10D.

```
Hmw4com   YSIAAPENEA INLGDIFAKG GNINVRAATI RNKGKLSADS VSKDKSGNIV
Hmw1com   YSIAAPENEA VNLGDIFAKG GNINVRAATI RNKGKLSADS VSKDKSGNIV
Hmw2com   YSIAAPENEA VNLGDIFAKG GNINVRAATI RNKGKLSADS VSKDKSGNIV

          301                                                  350
Hmw3com   LSAKEGEAEI GGVISAQNQQ AKGGKLMITG DKVTLKTGAV IDLSGKEGGE
Hmw4com   LSAKEGEAEI GGVISAQNQQ AKGGKLMITG DKVTLKTGAV IDLSGKEGGE
Hmw1com   LSAKEGEAEI GGVISAQNQQ AKGGKLMITG DKVTLKTGAV IDLSGKEGGE
Hmw2com   LSAKEGEAEI GGVISAQNQQ AKGGKLMITG DKVTLKTGAV IDLSGKEGGE

          351                                                  400
Hmw3com   TYLGGDERGE GKNGIQLAKK TTLEKGSTIN VSGKEKGGRA IVWGDIALID
Hmw4com   TYLGGDERGE GKNGIQLAKK TTLEKGSTIN VSGKEKGGRA IVWGDIALID
Hmw1com   TYLGGDERGE GKNGIQLAKK TTLEKGSTIN VSGKEKGGRA IVWGDIALID
Hmw2com   TYLGGDERGE GKNGIQLAKK TTLEKGSTIN VSGKEKGGRA IVWGDIALID
```

FIG.10E.

```
                401                                                 450
Hmw3com   GNINAQGK.D IAKTGGFVET SGHYLSIDDN AIVKTKEWLL DPENVTIEAP
Hmw4com   GNINAQGS.D IAKTGGFVET SGHDLSIGDD VIVDAKEWLL DPDDVSIETL
Hmw1com   GNINAQGSGD IAKTGGFVET SGHDLFIKDN AIVDAKEWLL DPDNVTINAE
Hmw2com   GNINAQGSGD IAKTGGFVET SGHYLSIESN AIVKTKEWLL DPDDVTIEAE

                451                                                 500
Hmw3com   SASRVELGAD RNSHSAEVIK VTLKKNNTSL TTLTNTTISN LLKSAHVVNI
Hmw4com   TSGRNNTGEN QGYTTGDGTK ESPKGNSISK PTLTNSTLEQ ILRRGSYVNI
Hmw1com   TAGRSNTSED DEYTGSGNSA STPKRNKE.K TTLTNTTLES ILKKGTFVNI
Hmw2com   DPLRNNTGIN DEFPTGTGEA SDPKKNSELK TTLTNTTISN YLKNAWTMNI

                501                                                 550
Hmw3com   TARRKLTVNS SISIERGSHL ILHSEGQGGQ GVQIDKDITS .E...GGNLT
Hmw4com   TANNRIYVNS SINLSNGS.L TLHTK...RD GVKINGDITS NE...NGNLT
Hmw1com   TANQRIYVNS SINL.SNGSL TLWSEGRSGG GVEINNDITT GDDTRGANLT
Hmw2com   TASRKLTVNS SINGSNGSHL ILHSKGQRGG GVQIDGDIT. ...SKGGNLT
```

FIG.10F.

```
                551
Hmw3com   IYSGGWVDVH KNITLGS.GF LNITTKEGDI AFEDKSGR..  ..NNLTITAQ
Hmw4com   IKAGSWVDVH KNITLGT.GF LNIVAGDS.V AFEREGDKAR NATDAQITAQ
Hmw1com   IYSGGWVDVH KNISLGAQGN INITAKQD.I AFEKGSNQV.  .....ITGQ
Hmw2com   IYSGGWVDVH KNITLD.QGF LNITA.AS.V AFEGGNNKAR DANNLTITAQ

                601
Hmw3com   GTITSG.NSN GFRFNNVSLN SLGGKLSFTD SREDRGRRTK GNISNKFDGT
Hmw4com   GTITVNKDDK QFRFNNVSIN GTGKGLKFIA NQN.......  .NFTHKFDGE
Hmw1com   GTIT.SGNQK GFRFNNVSLN GTGSGLQFTT KRTN.....K YAITNKFEGT
Hmw2com   GTVTITGEGK DFRANNVSLN GTGKGLNIIS SVNN......  ..LTHNLSGT

                651
Hmw3com   LNISGTVDIS MKAPKVSWFY RD.KGRTYWN VTTLNVTSGS KFNLSIDSTG
Hmw4com   INISGIVTIN QTTKKDVKYW NA.SKDSYWN VSSLTLNTVQ KFTF.IKFVD
Hmw1com   LNISGKVNIS MVLPKNESGY DKFKGRTYWN LTSLNVSESG EFNLTIDSRG
```

FIG.10G.

```
Hmw2com  INISGNITIN QTTRKNTSYW QTSHD.SHWN VSALNLETGA NFTF.IKYIS

              701                                              750
Hmw3com  SGSTG...PS IRNA..ELNG ITFN....KA TFNIAQGSTA NFSIKASIMP
Hmw4com  SGSNS...QD LRSSRRSFAG VHFNGIGGKT NFNIGANAKA LFKLKPNAAT
Hmw1com  SDSAGTLTQ. ....PYNLNG ISFN...KDT TFNVERNARV NFDIKAPIGI
Hmw2com  SNSKGLTTQY RSSAGVNFNG V..N...GNM SFNLKEGAKV NFKLKPNENM

              751                                              800
Hmw3com  FKSNANYAL. FNEDISVSG. .GGSVNFKLN ASSSNIQTPG VIIKSQNFNV
Hmw4com  DPKKELPIT. FNANITATGN SDSSVMFDIH A...NLTSRA AGINMDSINI
Hmw1com  NKYSSLNYAS FNGNISVSG. .GGSVDFTLL ASSSNVQTPG VVINSKYFNV
Hmw2com  NTSKPLPI.R FLANITATG. .GGSVFFDIY ANHS...GRG AELKMSEINI

              801                                              850
Hmw3com  SGGSTLNLKA EGSTETAFSI ENDLNLNATG GNITIRQVEG T..DSRVNKG
Hmw4com  TGGLDFSITS HNRNSNAFEI KKDLTINATG SNFSLKQTKD SFYNEYSKHA
```

FIG.10H.

```
Hmw1com   STGSSLRFKT SGSTKTGFSI EKDLTLNATG GNITLLQVEG T..DGMIGKG
Hmw2com   SNGANFTLNS HVRGDDAFKI NKDLTINATN SNFSLRQTKD DFYDGYARNA

                851                                           900
Hmw3com   VAAKKNITFK GGNITFGSQK ATTEIKGNVT INKNTNATLR GANFAEN...
Hmw4com   INSSHNLTIL GGNVTLGGEN SSSSITGNIN ITNKANVTLQ ADTSNSNTGL
Hmw1com   IVAKKNITFE GGNITFGSRK AVTEIEGNVT INNNANVTLI GSDFDNHQ..
Hmw2com   INSTYNISIL GGNVTLGGQN SSSSITGNIT IEKAANVTLE ANNAPNQQNI

                901                                           950
Hmw3com   KSPLNIAGNV INNGNLTTAG SIINIAGNLT VSKGANLQAI TNYTFNVAGS
Hmw4com   KKRTLTLGNI SVEGNLSLTG ANANIVGNLS IAEDSTFKGE ASDNLNITGT
Hmw1com   KPLTIKKDVI INSGNLTAGG NIVNIAGNLT VESNANFKAI TNFTFNVGGL
Hmw2com   RDRVIKLGSL LVNGSLSLTG ENADIKGNLT ISESATFKGK TRDTLNITGN

                951                                          1000
```

FIG.10I.

```
Hmw3com   FDNNGASNIS IARGGAKFK. DINNTSSLNI TTNSDTTYRT IIKGNISNKS
Hmw4com   FTNNGTANIN IKQGVVKLQG DINNKGGLNI TTNASGTQKT IINGNITNEK
Hmw1com   FDNKGNSNIS IAKGGARFK. DIDNSKNLSI TTNSSSTYRT IISGNITNKN
Hmw2com   FTNNGTAEIN ITQGVVKLG. NVTNDGDLNI TTHAKRNQRS IIGGDIINNK

          1001                                              1050
Hmw3com   GDLNIIDKKS DAEIQIGGNI SQKEGNLTIS SDKVNITNQI TIKAGVEGGR
Hmw4com   GDLNIKNIKA DAEIQIGGNI SQKEGNLTIS SDKVNITNQI TIKAGVEGGR
Hmw1com   GDLNITNEGS DTEMQIGGDI SQKEGNLTIS SDKINITKQI TIKAGVDGEN
Hmw2com   GSLNITDSNN DAEIQIGGNI SQKEGNLTIS SDKINITKQI TIKKGIDGED

          1051                                              1100
Hmw3com   SDSSEAENAN LTIQTKELKL AGDLNISGFN KAEITAKNGS DLTIGNASGG
Hmw4com   SDSSEAENAN LTIQTKELKL AGDLNISGFN KAEITAKNGS DLTIGNASGG
Hmw1com   SDSDATNNAN LTIKTKELKL TQDLNISGFN KAEITAKDGS DLTIGNTNSA
Hmw2com   SSSDATSNAN LTIKTKELKL TEDLSISGFN KAEITAKDGR DLTIGNSNDG
```

FIG.10J.

```
                1101                                                1150
Hmw3com   N..ADAKKVT FDKVKDSKIS TDGHNVTLNS EVKT..SNGS SNAGNDNSTG
Hmw4com   N..ADAKKVT FDKVKDSKIS TDGHNVTLNS EVKT..SNGS SNAGNDNSTG
Hmw1com   D.GTNAKKVT FNQVKDSKIS ADGHKVTLHS KVETSGSNNN TEDSSDNNAG
Hmw2com   NSGAEAKKVT FNNVKDSKIS ADGHNVTLNS KVKTSSSNGG RESNSDNDTG

                1151                                                1200
Hmw3com   LTISAKDVTV NNNVTSHKTI NISAAAGNVT TKEGTTINAT TGSVEVTAQN
Hmw4com   LTISAKDVTV NNNVTSHKTI NISAAAGNVT TKEGTTINAT TGSVEVTAQN
Hmw1com   LTIDAKNVTV NNNITSHKAV SISATSGEIT TKTGTTINAT TGNVEIT...
Hmw2com   LTITAKNVEV NKDVTSLKTV NITA.SEKVT TTAGSTINAT NGKASIT...

                1201                                                1250
Hmw3com   GTIKGNITSQ NVTVTATENL VTTENAVINA TSGTVNISTK TGDIKGGIES
Hmw4com   GTIKGNITSQ NVTVTATENL VTTENAVINA TSGTVNISTK TGDIKGGIES
Hmw1com   .......... .......... .......... ........AQ TGDIKGGIES
```

FIG.10K.

```
Hmw2com   ..... ..... ..... .....TK T.....

          1251                                                 1300
Hmw3com   TSGNVNITAS GNTLKVSNIT GQDVTVTADA GALTTTAGST ISATTGNANI
Hmw4com   TSGNVNITAS GNTLKVSNIT GQDVTVTADA GALTTTAGST ISATTGNANI
Hmw1com   SSGSVTLTAT EGALAVSNIS GNTVTVTANS GALTTLAGST IKG.TESVTT
Hmw2com   ..... ..... ..... ..... .....

          1301                                                 1350
Hmw3com   TTKTGDINGK VESSSGSVTL VATGATLAVG NISGNTVTIT ADSGKLTSTV
Hmw4com   TTKTGDINGK VESSSGSVTL VATGATLAVG NISGNTVTIT ADSGKLTSTV
Hmw1com   SSQSGDIG.. ..... .....G TISGGTVEVK ATESLTTQSN
Hmw2com   ....GDIS.. ..... .....G TISGNTVSVS ATVDLTTKSG

          1351                                                 1400
Hmw3com   GSTINGTNSV TTSSQSGDIE GTISGNTVNV TASTGDLTIG NSAKVEAKNG
Hmw4com   GSTINGTNSV TTSSQSGDIE GTISGNTVNV TASTGDLTIG NSAKVEAKNG
```

FIG.10L.

```
Hmw1com   SKIKATTGEA NVTSATGTIG GTISGNTVNV TANAGDLTVG NGAEINATEG
Hmw2com   SKIEAKSGEA NVTSATGTIG GTISGNTVNV TANAGDLTVG NGAEINATEG

                1401                                          1450
Hmw3com   AATLTAESGK LTTQTGSSIT SSNGQTTLTA KDSSIAGNIN AANVTLNTTG
Hmw4com   AATLTAESGK LTTQTGSSIT SSNGQTTLTA KDSSIAGNIN AANVTLNTTG
Hmw1com   AATLTTSSGK LTTEASSHIT SAKGQVNLSA QDSSVAGSIN AANVTLNTTG
Hmw2com   AATLTATGNT LTTEAGSSIT STKGQVDLLA QNSSIAGNIN AANVTLNTTG

                1451                                          1500
Hmw3com   TLTTTGDSKI NATSGTLTIN AKDAKLDGAA SGDRTVVNAT NASGSGNVTA
Hmw4com   TLTTTGDSKI NATSGTLTIN AKDAKLDGAA SGDRTVVNAT NASGSGNVTA
Hmw1com   TLTTVKGSNI NATSGTLTIN AKDAELNGAA LGNHTVVNAT NANGSGSVIA
Hmw2com   TLTTVAGSDI KATSGTLTIN AKDAKLNGDA SGDSTEVNAV NASGSGSVTA

                1501                                          1550
```

FIG.10M.

```
Hmw3com  KTSSSVNITG DLNTINGLNI ISENGRNTVR LRGKEIDVKY IQPGVASVEE
Hmw4com  KTSSSVNITG DLNTINGLNI ISENGRNTVR LRGKEIDVKY IQPGVASVEE
Hmw1com  TTSSRVNITG DLITINGLNI ISKNGINTVL LKGVKIDVKY IQPGIASVDE
Hmw2com  ATSSSVNITG DLNTVNGLNI ISKDGRNTVR LRGKEIEVKY IQPGVASVEE

                1551
Hmw3com  VIEAKRVLEK VKDLSDEERE TLAKLGVSAV RFVEPNNAIT VNTQNEFTTK
                                                              1600
Hmw4com  VIEAKRVLEK VKDLSDEERE TLAKLGVSAV RFVEPNNAIT VNTQNEFTTK
Hmw1com  VIEAKRILEK VKDLSDEERE ALAKLGVSAV RFIEPNNTIT VDTQNEFATR
Hmw2com  VIEAKRVLEK VKDLSDEERE TLAKLGVSAV RFVEPNNTIT VNTQNEFTTR

                1601
                                                      1632
Hmw3com  PSSQVTISEG KACFSSGNGA RVCTNVADDG QQ
Hmw4com  PSSQVTISEG KACFSSGNGA RVCTNVADDG QQ
Hmw1com  PLSRIVISEG RACFSNSDGA TVCVNIADNG R.
Hmw2com  PSSQVIISEG KACFSSGNGA RVCTNVADDG QP
```

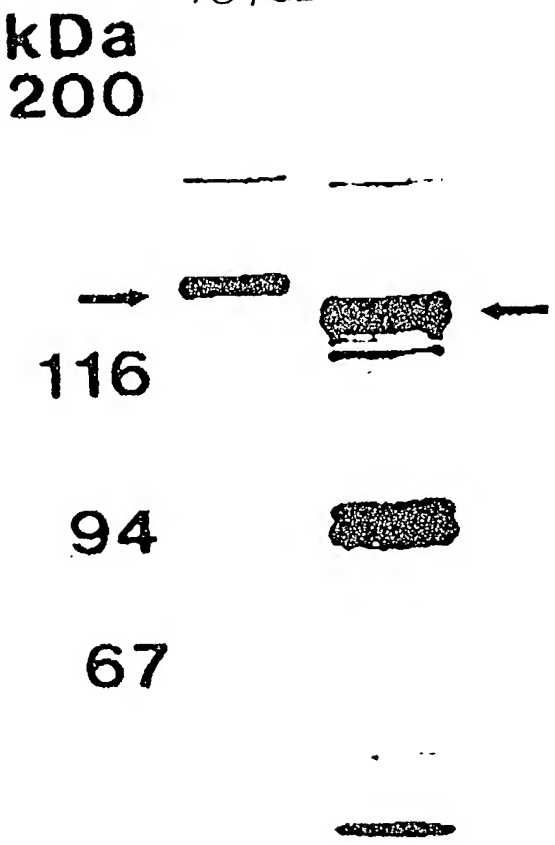


FIG. 11

WESTERN IMMUNOBLOT ASSAY OF PHAGE LYSATES CONTAINING EITHER THE HMW1 OR HMW2 RECOMBINANT PROTEINS. LYSATES WERE PROBED WITH AN *E. COLI*-ABSORBED ADULT SERUM SAMPLE WITH HIGH-TITER ANTIBODY AGAINST HIGH-MOLECULAR-WEIGHT PROTEINS. THE ARROWS INDICATE THE MAJOR IMMUNOREACTIVE PROTEIN BANDS OF 125 AND 120 kDa IN THE HMW1 AND HMW2 LYSATES, RESPECTIVELY.

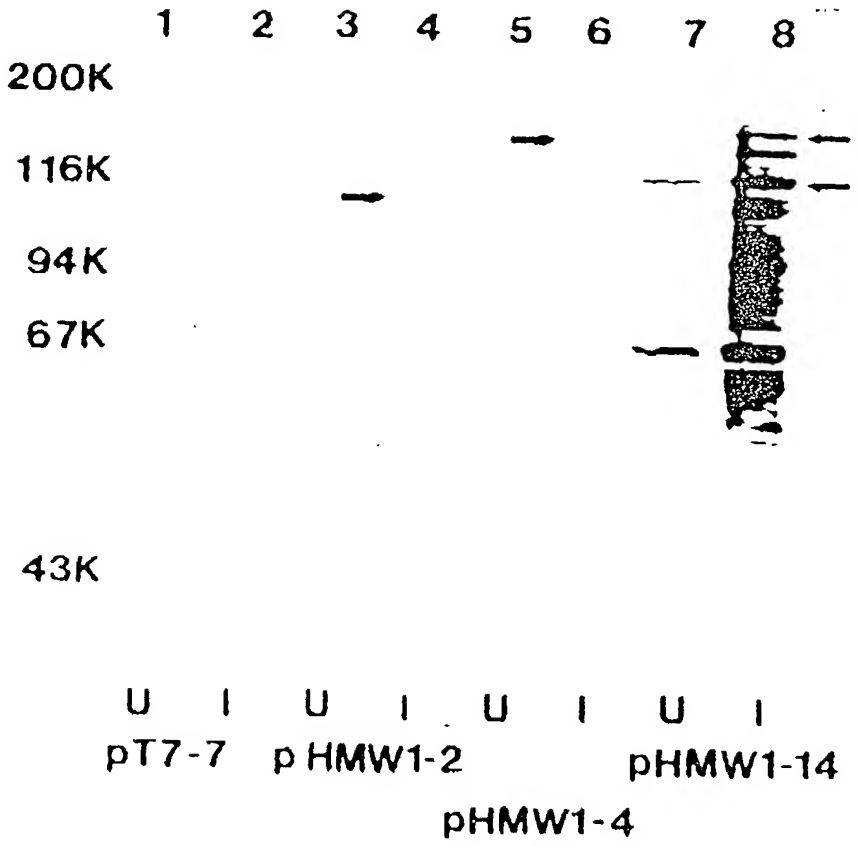


FIG. 12

WESTERN IMMUNOBLOT ASSAY OF CELL SONICATES PREPARED FROM *E. COLI* TRANSFORMED WITH PLASMID pT7-7 (LANES 1 AND 2) pHMW1-2 (LANES 3 AND 4), pHMW1-4 (LANES 5 AND 6), OR pHMW1-14 (LANES 7 AND 8). THE SONICATES WERE PROBED WITH AN *E. COLI*-ABSORBED ADULT SERUM SAMPLE WITH HIGH -TITER ANTIBODY AGAINST HIGH - MOLECULAR -WEIGHT PROTEINS. LANES LABELED U AND I REPRESENT SONICATES PREPARED BEFORE AND AFTER INDUCTION OF THE GROWING SAMPLES WITH IPTG, RESPECTIVELY. THE ARROWS INDICATE PROTEIN BANDS OF INTEREST AS DESCRIBED IN THE TEXT.

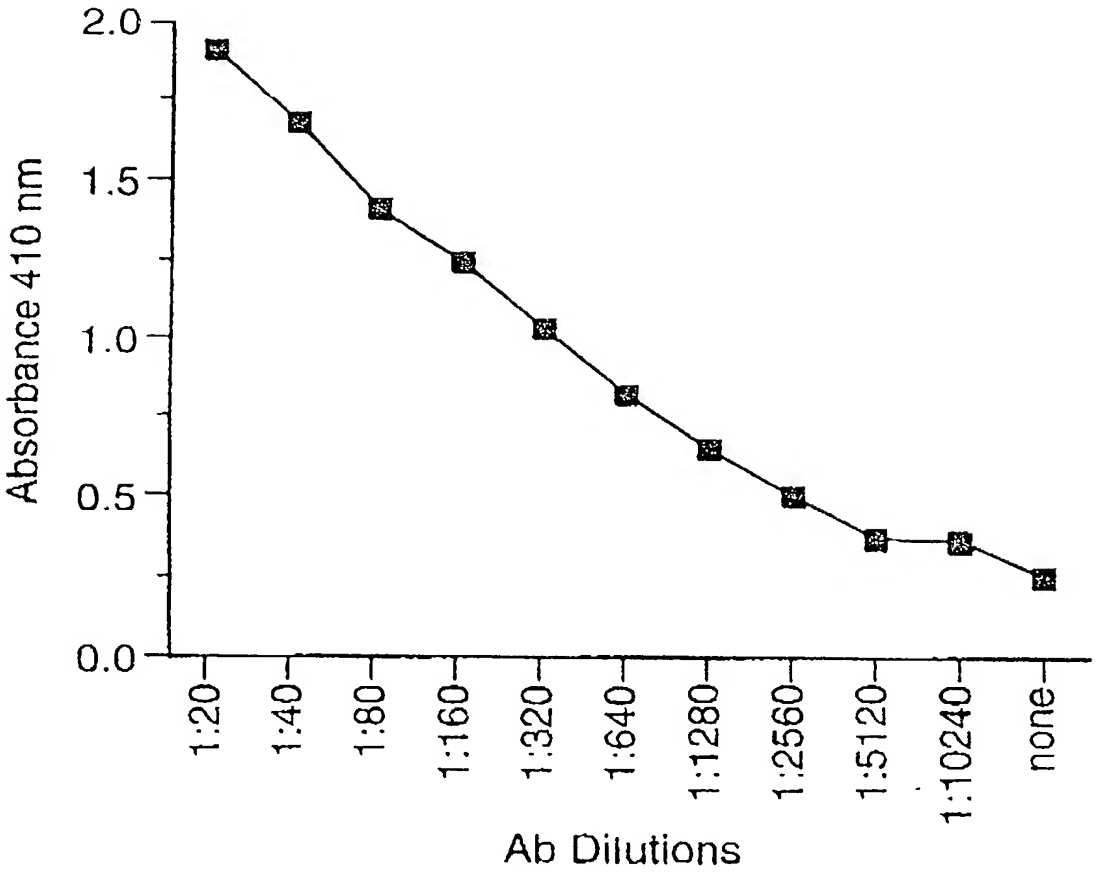


FIG. 13

ELISA WITH rHMW1 ANTISERUM ASSAYED AGAINST PURIFIED FILAMENTOUS HEMAGLUTININ OF B. *PERTUSSIS*. Ab, ANTIBODY.

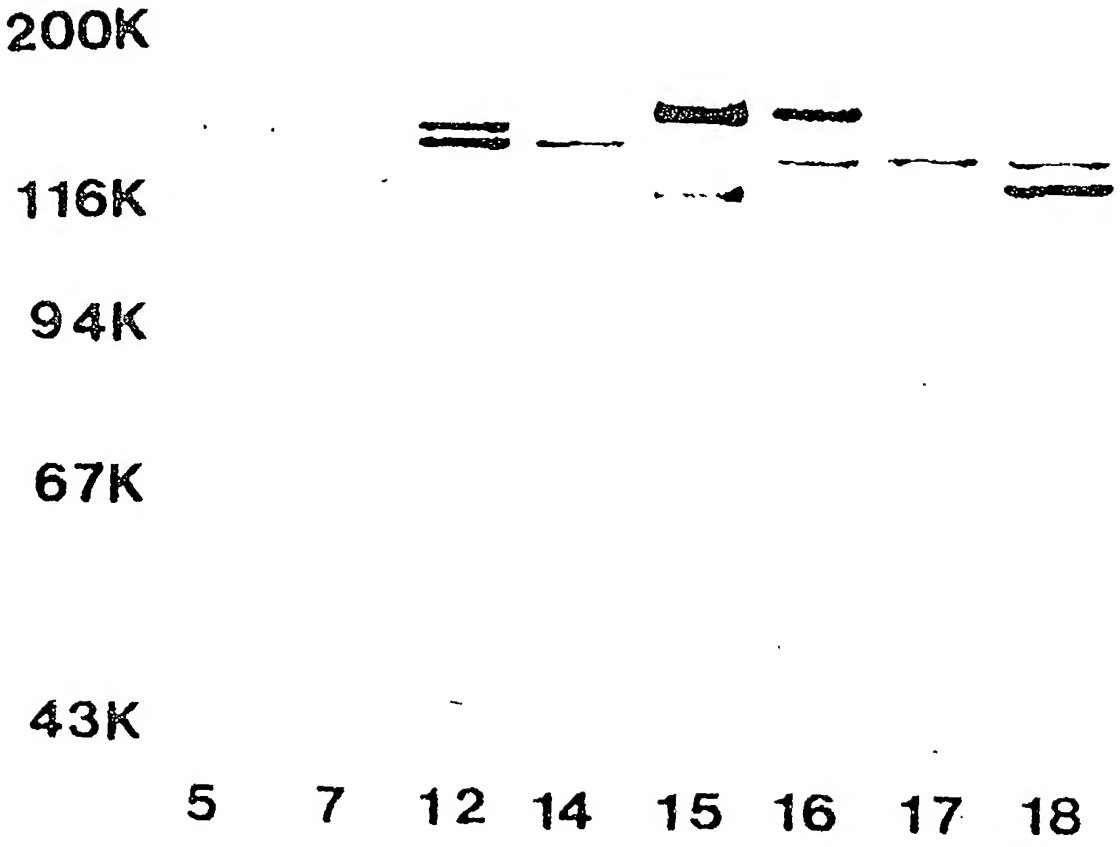


FIG. 14

WESTERN IMMUNOBLOT ASSAY OF CELL SONICATES FROM A PANEL OF EPIDEMIOLOGICALLY UNRELATED NONTYPEABLE *H. INFLUENZAE* STRAINS.THE SONICATES WERE PROBED WITH RABBIT ANTISERUM PREPARED AGAINST HMW1-4 RECOMBINANT PROTEIN. THE STRAIN DESIGNATIONS ARE INDICATED BY THE NUMBERS BELOW EACH LANE.

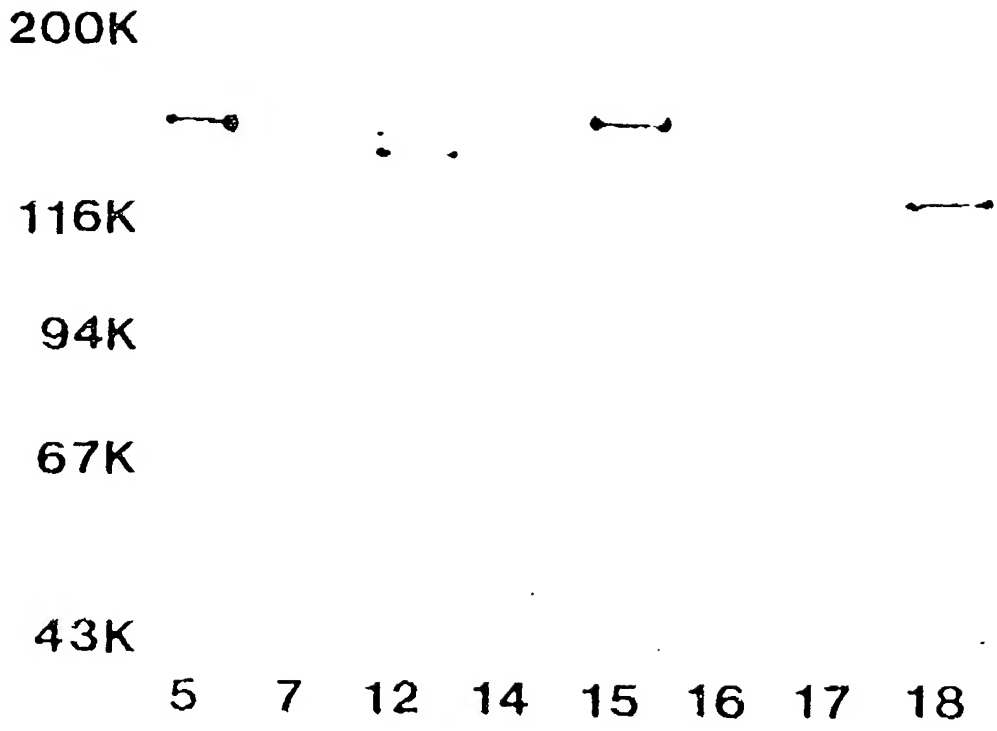


FIG. 15

WESTERN IMMUNOBLOT ASSAY OF CELL SONICATES FROM A PANEL OF EPIDEMIOLOGICALLY UNRELATED NONTYPEABLE H. *INFLUENZAE* STRAINS.THE SONICATES WERE PROBED WITH MONOCLONAL ANTIBODY X3C, A MURINE IgG ANTIBODY WHICH RECOGNIZES THE FILAMENTOUS HEMAGGLUTININ OF *B. PERTUSSIS* (13).THE STRAIN DESIGNATIONS ARE INDICATED BY THE NUMBERS BELOW EACH LANE.

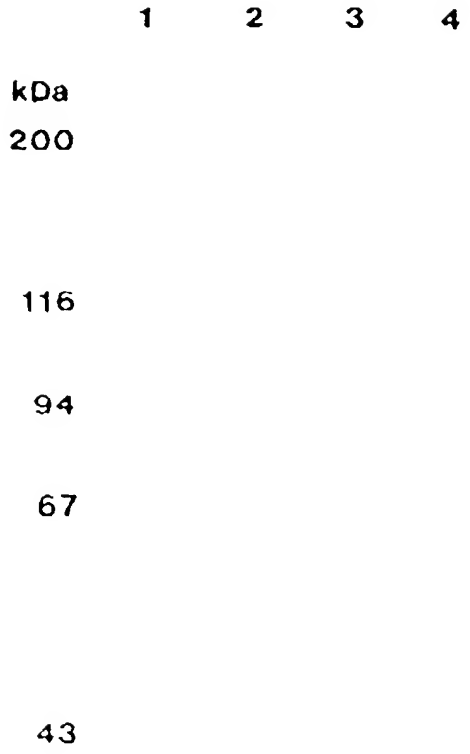
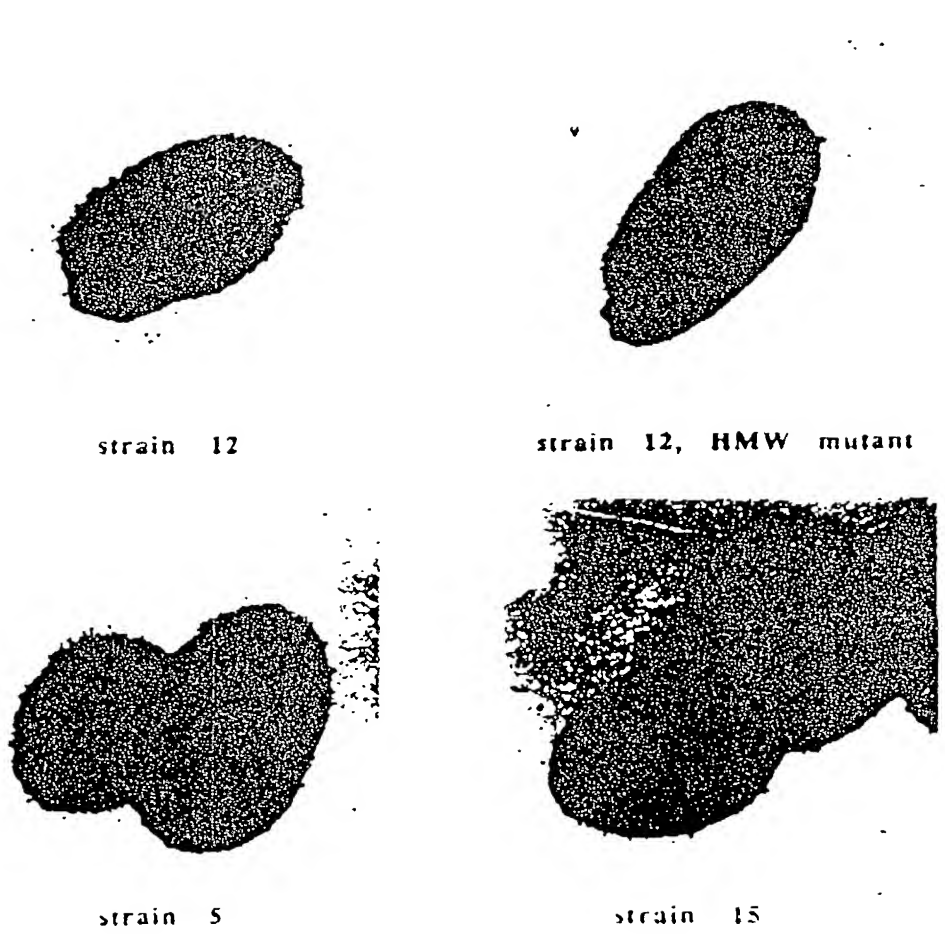


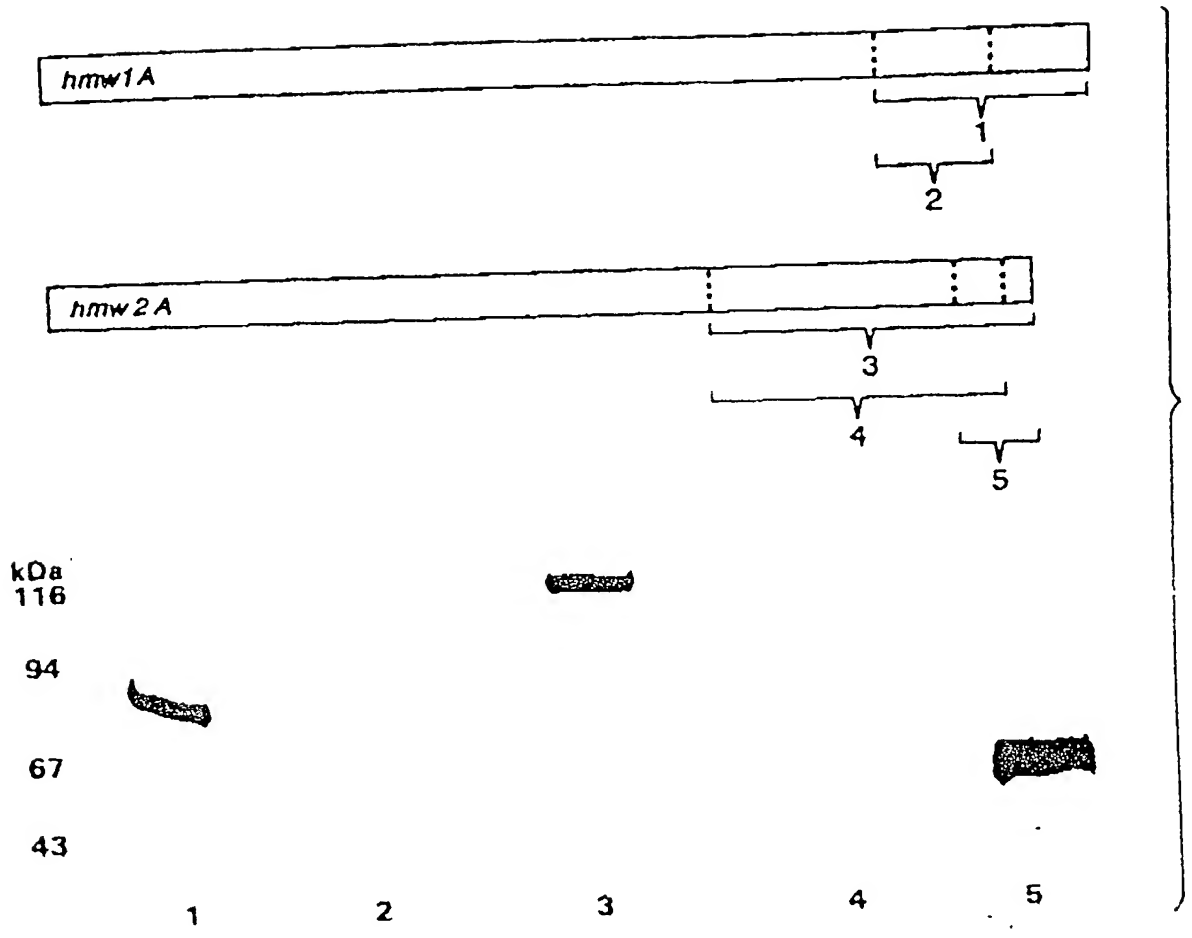
FIG. 16

IMMUNOBLOT ASSAY OF CELL SONICATES OF NONTYPABLE *H. INFLUENZAE* STRAIN 12 DERIVATIVES. THE SONICATES WERE PROBED WITH RABBIT ANTISERUM PREPARED AGAINST HMW-1 RECOMBINANT PROTEIN. LANES: 1, WILD-TYPE STRAIN; 2, HMW-2$^-$ MUTANT; 3, HMW-1$^-$ MUTANT; 4, HMW-1$^-$ / HMW-2$^-$ DOUBLE MUTANT.



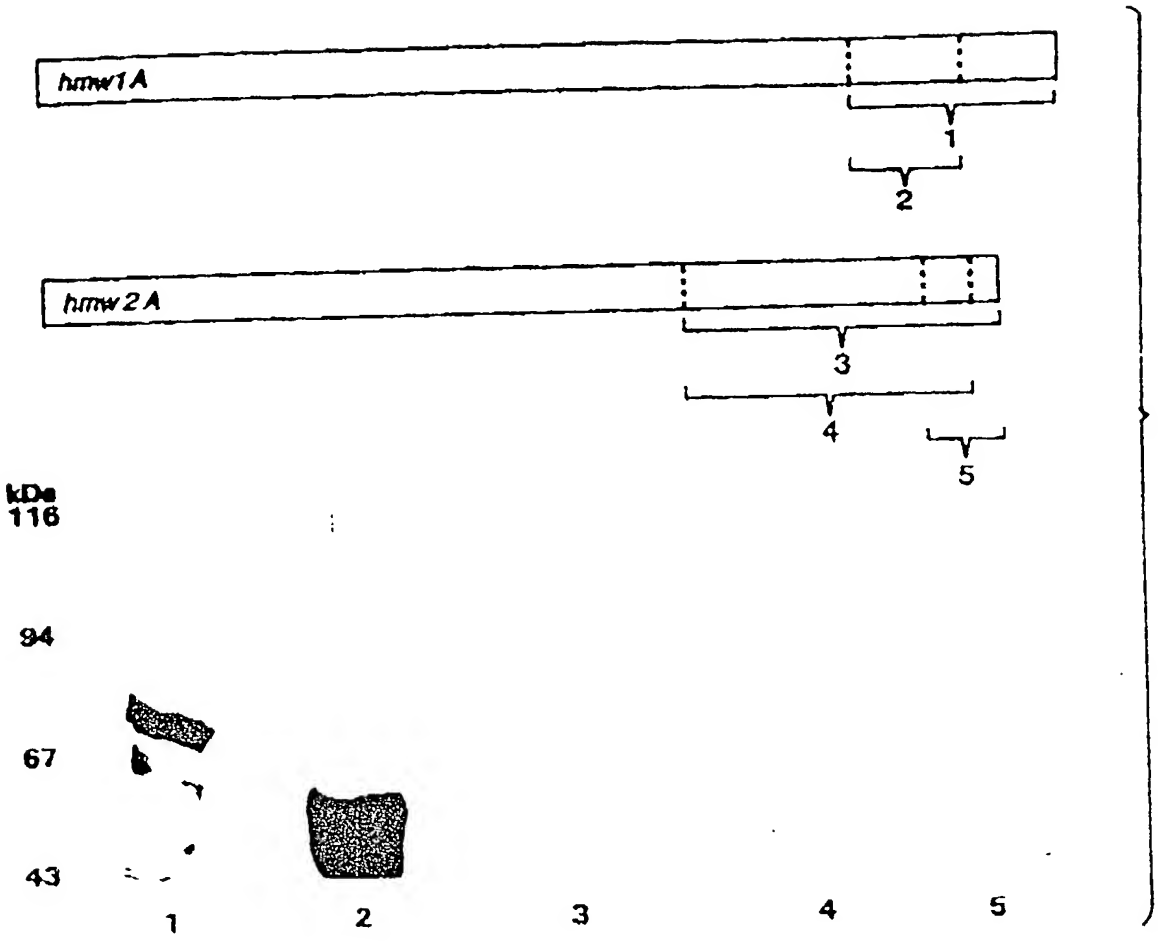
IMMUNOELECTRON MICROSCOPY WITH Mab AD6

FIG.20



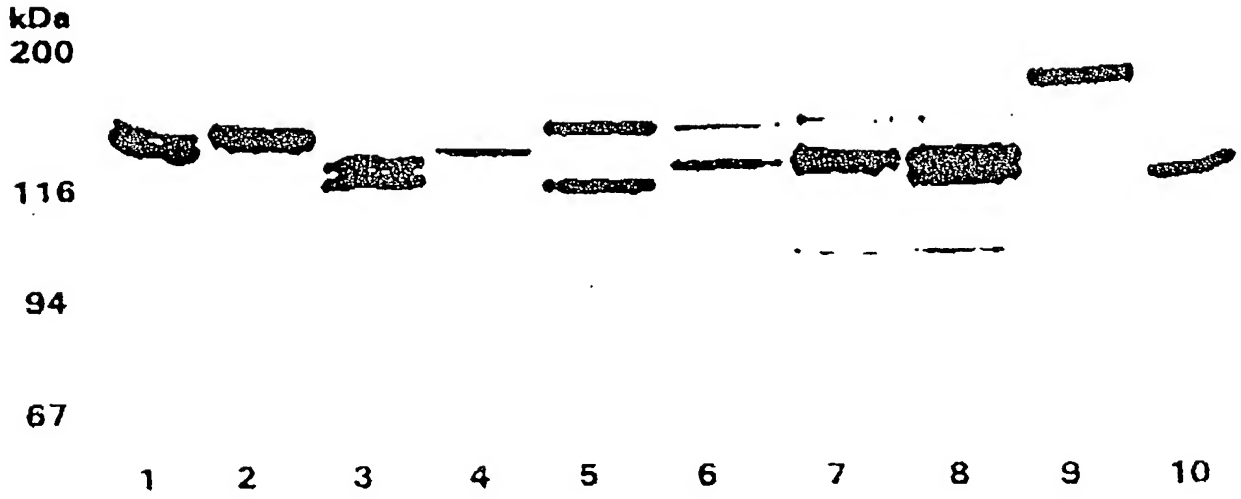
WESTERN IMMUNOBLOT ASSAY WITH Mab AD6 AND HMW1A OR HMW2A RECOMBINANT PROTEINS

FIG.21



WESTERN IMMUNOBLOT ASSAY WITH Mab 10C5 AND HMW1A OR HMW2A RECOMBINANT PROTEINS

FIG.22



WESTERN IMMUNOBLOT ASSAY WITH Mab AD6 AND TEN UNRELATED NONTYPABLE *HAEMOPHILUS INFLUENZAE*

FIG.23